2008

Context visuals in L2 listening tests: the effectiveness of photographs and video vs. audio-only format

Ruslan S. Suvorov
Iowa State University

Context visuals in L2 listening tests:

The effectiveness of photographs and video vs. audio-only format

by

Ruslan S Suvorov

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Major: Teaching English as a Second Language/Applied Linguistics (Language Assessment) and (English for Specific Purposes)

Program of Study Committee:
Volker Hegelheimer, Major Professor
Dan Douglas
Charles Kostelnick

Iowa State University

Ames, Iowa

2008

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Although visual support in the form of pictures and video has been widely used in language teaching, there appears to be a dearth of research on the role of visual aids in L2 listening tests (Buck, 2000; Ockey, 2007) and the absence of sound theoretical perspectives on this issue (Ginther, 2001; Gruba, 1999). The existing studies of the role of visual support in L2 listening tests yielded inconclusive results. While some studies showed that visuals can improve test-takers' performance on L2 listening tests (e.g., Ginther, 2002), others revealed no facilitative effect of visuals on listening comprehension of test-takers (e.g., Coniam, 2001; Gruba, 1993; Ockey, 2007).

The given study, conducted at Iowa State University in Spring 2008, investigated the influence of context visuals, namely a single photograph and video, on test-takers' performance on a computer-based Listening Test developed specifically for this study. The Listening Test, consisting of six listening passages and 30 multiple-choice questions, was administered to 34 international students from three English listening classes. In particular, the study examined whether test-takers perform differently on three types of listening passages: passages with a single photograph, video-mediated listening passages, and audio-only listening passages. In addition, participants' responses on the Post-Test Questionnaire were analyzed to determine whether their preferences of visual stimuli in listening tests corresponded with their actual performance on different types of visuals.

The results indicated that while no difference was found between the scores for photo-mediated and audio-only listening passages, participants' performance on video-mediated listening passages was significantly lower.

CHAPTER 1. INTRODUCTION

This thesis is concerned with the role of visual support in second language (L2) listening comprehension. Specifically, this study focuses on the use of a single photograph and video in L2 listening tests and the impact of these visual elements in terms of their facilitative or distracting effect on L2 test-takers' performance. Although visuals have been used in L2 teaching and testing for a number of decades (Coniam, 2001; Ginther, 2001, 2002; Ockey, 2007), there is insufficient empirical evidence to date concerning the role of visual support in assessing L2 learners' listening comprehension.

Statement of the Problem

In light of advances in computer-assisted language learning (CALL) and the use of technology in testing, listening tests (such as the ones that are included in the listening section of TOEFL iBT) are being offered both online and in the offline medium. Although changes in technology have fueled the interest in visual instructional materials (Wetzel, Radtke, & Stern, 1994), there appears to be a dearth of research on the role of visual aids in L2 listening tests (Buck, 2000; Ockey, 2007) and the absence of sound theoretical perspectives on this issue (Ginther, 2001; Gruba, 1999).

Early research on visual support suggested that one way to promote L2 listening comprehension was by using pictures (e.g., Brasnford & Johnson, 1972, cited in Chung, 1994; Mueller, 1980). With the advent of accessible video equipment in the late 1970s and early 1980s, research studies on the use and the role of video for the development of L2 listening skills started to emerge as well (Gruba, 1999).

Researchers tend to agree that compared to an audio medium, video is more authentic in terms of context, discourse, paralinguistic features, and culture (e.g., Coniam, 2001). However, in spite of the putative merits of video, there is a dearth of research comparing the effects of audio and video modes in listening tests (Buck, 2001; Chung, 1994; Coniam, 2001).

The existing studies of the role of visual support both in L2 listening comprehension and L2 listening tests yield inconclusive results. Research on the use of visuals in listening comprehension suggests that the use of video can promote students' listening skills (Secules, Herron, & Tomasello, 1992) and is more helpful for less proficient language learners, especially when the latter encounter difficult texts (Mueller, 1980; Rubin, 1995). In other studies, it was found that non-verbal clues have little, if any, effect on facilitating the understanding of verbal information (Lynch, 1998). Thus, many experts in L2 listening call for more research to better understand the role of visual information in listening comprehension (Buck, 2001; Rubin, 1995).

The results of research that focuses on the role of visuals in L2 listening tests appear to be inconclusive as well. While some studies showed that visuals can improve students' performance on listening tests (e.g., Ginther, 2002), others evinced no facilitative effect, or in some cases even detrimental effect, of visuals on test-takers' listening comprehension (e.g., Coniam, 2001; Gruba, 1993; Ockey, 2007). Therefore, further research is needed to determine the role of visual support in L2 listening tests.

Aim and Scope of the Study

The aim of this thesis is to investigate the role of visual support, such as a single photograph and video, in second language listening tests. In particular, the study examines

whether there exists a difference in students' performance on three types of listening tests: the listening tests that are provided together with photographs, the listening tests that contain video, and the listening tests without any visual aids.

To achieve this aim, a computer-based Listening Test (LT) consisting of six listening passages and 30 multiple-choice questions was developed and administered to 34 international students enrolled in three ESL listening classes at a large public university in the Midwest of the USA. To enhance the validity and reliability of the Listening Test, a pilot study was conducted and an expert analysis was employed to assess the listening passages and the question items used in the Listening Test. The process of data collection for the main study was divided into three stages, with each stage being devoted to gathering data in one of the three listening classes. Besides the Listening Test, a Pre-Test Questionnaire, a Perceptual Learning Preferences Survey, and a Post-Test Questionnaire were developed and used for data collection purposes. The data gathered in the study underwent quantitative analysis, and conclusions were drawn on the basis of the obtained results.

Research Questions

This study will address the following three research questions.

Research Question #1: Is there a difference among different types of visual input, – namely a single photograph, video, and audio-only format, – in an L2 listening test in terms of their effect on L2 test-takers' performance?

Research Question #2: Does the use of visuals, namely a single photograph and video, in an L2 listening test facilitate test-takers' performance?

Research Question #3: Do test-takers' preferences of visual stimuli in listening tests correspond to their actual performance on different types of visuals in the Listening Test?

Structure of the Study

This thesis is organized into five chapters. The purpose of Chapter 2 is to provide theoretical perspectives and recent research findings on the use of visuals in listening comprehension and listening tests. Chapter 3 describes the methods used for collecting and analyzing the data for this study, including participants, materials, setting, and procedures. Chapter 4 focuses on the results of the quantitative analyses of the collected data and their discussion. Finally, Chapter 5 concludes with implications, limitations of the study, and ideas for future research.

CHAPTER 2. LITERATURE REVIEW

The aim of this chapter is to establish a theoretical background for the current study by reviewing the existing studies and theories pertaining to first and second language listening comprehension and testing of listening skills. Chapter 2 consists of seven sections. The first section defines listening comprehension and discusses some factors that affect listening. The second section describes the existing models of listening comprehension. In the third section, several classifications of visuals are given and their effects on learning are discussed. The fourth section analyzes multimedia language learning. The fifth section presents research on the use of visuals in L2 listening comprehension, while the sixth section provides an overview of research on the use of visuals in listening tests. Finally, the last section discusses some issues pertaining to the construct definition of L2 listening ability.

Views of Listening Comprehension

Listening is a key language skill that is important for the language acquisition process (Brett, 1997; Rubin, 1995). However, sound is not the only means of conveying information in spoken discourse. In real-life communication, the verbal information is often accompanied by visual information. Therefore, as stated by Buck (2001), "the common practice of playing a disembodied recording from an audio-player does not create a very realistic listening situation" (p. 253).

There exist a number of various definitions of the listening comprehension process, all of them varying to some degree. According to one of the earlier definitions given by Lado (1961), listening comprehension can be defined as "recognition control of the signaling elements of the language in communication situations" (p. 206). Later definitions of listening

comprehension specify the types of stimuli or information exchanged between the speaker and the listener. For example, Coakley and Wolvin (1986) claim that listening is "a complex communication behavior, involving a process of receiving, attending to, and assigning meaning to verbal and/or non-verbal stimuli" (p. 20), while Rubin (1995) states that listening comprehension "consists of processing information which listeners get from visual and auditory clues in order to define what is going on and what the speakers are trying to express" (p. 151). Mueller (1980) speculated that listening comprehension results from "a complex interplay of linguistic and extralinguistic, contextual (often visual) information cues" (p. 335).

Thus, listening can be defined as a communication activity, in which the listener receives a message through the auditory, visual, and attention processors. Chung (1994) claims that listening comprehension data can be divided into two parts: the message sent to the listener and the message understood by the listener. The message sent consists of three types of information: oral (words and sentences), paralinguistic (extra-speech sounds, such as hissing and whistling, as well as quality of voice, voice pitch, and rate of speech), and visual (images, gestures, facial expressions, etc.).

Coakley and Wolvin (1986) developed a taxonomy of listening functions: discriminative, comprehensive, therapeutic, critical, and appreciative. Discriminative listening involves the development of "careful concentration and sensitivity to the various stimuli in order to interpret them meaningfully" (Coakley & Wolvin, 1986, p. 18). The purpose of comprehensive listening is to understand a message. Therapeutic listening implies putting the listener in the position of the speaker to understand the feelings and thoughts of the latter and to provide a supportive communication. Critical listening involves a critical

evaluation of the message by the listener. Finally, appreciative listening deals with sensual enjoyment and appreciation of the aural message conveyed to the listener.

Many researchers (e.g., Gruba, 1997; Ockey, 2007) maintain that the process of listening comprehension can be affected by a number of different factors, such as rate of speech, prosody, accent, phonology, hesitations, background knowledge, and rhetorical signaling cues. A number of researchers argue that lip movements of the speaker provide information and help the listener better understand what is being said (Kellerman, 1990, cited in Buck, 2001; Ockey, 2007). Furthermore, body movement, gestures, and facial expressions have also been mentioned as non-verbal signals that complement the verbal information (Altman, 1990, cited in Coniam, 2001; Buck, 2001; Ockey, 2007; Rubin, 1995).

Baltova (1994) argues that "in real-life listening comprehension we not only "listen" but more often than not "view" the message as well, and interpret the two modes of information in a similar way" (p. 508). She claims that this concept of "viewing comprehension," which was introduced by Riley (1979, cited in Baltova, 1994), is very important in communication and implies the processing of visual cues as opposed to "listening" per se, which involves the auditory perception and interpretation of verbal information. The process of listening that involves "making sense" of the received input also utilizes the listener's cultural and educational background knowledge (Rubin, 1995, p. 151). However, if the visual input does not fit into the listener's cultural expectations or background knowledge, it can be confusing and impede listening comprehension. Visuals can also be distracting or misleading when there is little or no relationship between what is said and what is shown (Rubin, 1995). Even though visual information does seem to play an

important role in oral communication, it is not clear exactly how listeners make use of various visual clues available in the process of communication.

Models of Listening Comprehension

Although there is no commonly accepted theory explaining the process of listening comprehension (Ockey, 2007), several models of listening comprehension have been proposed by different researchers (e.g., Brindley, 1998; Gruba, 1999).

One of the models of listening comprehension is based on an assumption that in order to comprehend auditory input, listeners can utilize a number of hierarchically arranged listening skills that vary from basic literal understanding of a fact to making inferences and the speaker's meaning (Brindley, 1998). Another model suggests that there is a "move away from the notion of listening as auditory discrimination and decoding of decontextualized utterances towards a much more complex and interactive model which reflects the ability to understand authentic discourse in context" (Brindley, 1998, p. 172). There is also a connectionist cognitive processing model proposed by Gruba (1999). This model is based on the idea that the processing of numerous incoming stimuli, including visual stimuli, in the human brain occurs simultaneously. However, connectionist cognitive processing model does not explain the impact visual stimuli might have on auditory comprehension.

The process of listening comprehension involves two stages: a listening stage and a response stage (Bejar, Douglas, Jamieson, Nissan, & Turner (2000). The listening stage includes the processing of an acoustic signal by receptive and cognitive processes, the activation of three types of knowledge (i.e., situational knowledge (SK), linguistic knowledge (LK), and background knowledge (BK)), and the transformation of the acoustic signal into a set of propositions (PR). On the basis of these sets of propositions, the language

learner produces a response, the adequacy of which is mediated by the learner's knowledge and cognitive factors (see Figure 2.1).

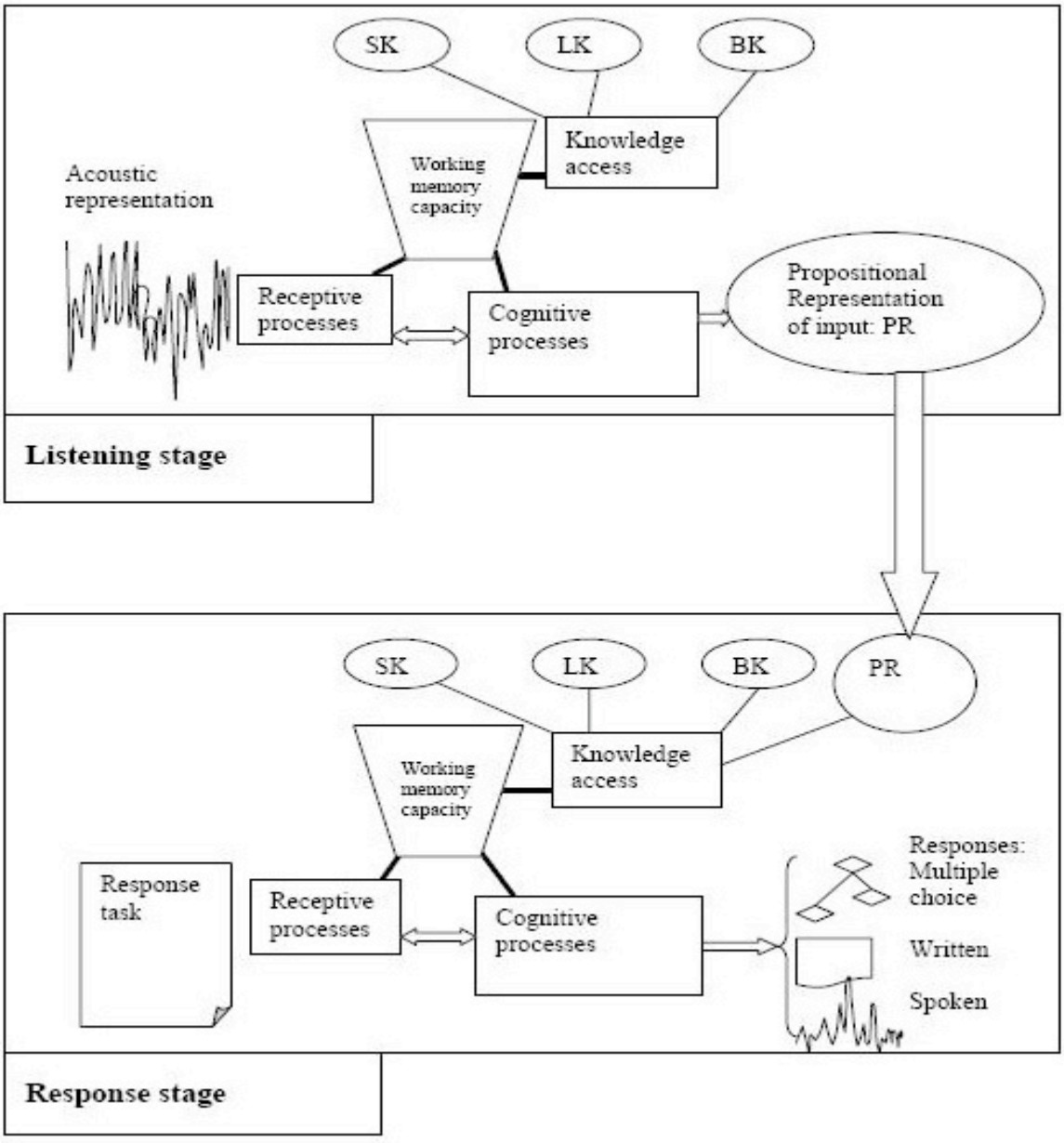

Figure 2.1. Listening stage and response stage of the listening process (adapted from Bejar et al., 2000, p. 3).

Thus, the existing models suggest that listening comprehension, which employs a hierarchy of listening skills, is a complex process that involves processing of auditory and

visual incoming stimuli in the human brain, activation of different types of knowledge, and production of a response.

Types of Visuals and Their Effects on Learning

As many researchers agree that visual elements are important for language learning, visuals are being widely used by language instructors. However, it should be noted that not all visuals are the same and that the use of different visuals can have either a facilitative or a debilitating effect that can lead to different learning outcomes.

Bejar et al. (2000) and Ginther (2002) differentiate between two types of visuals: context (or situation) visuals and content visuals. Context visuals are visuals that provide information about the context for the verbal exchanges, such as the participants, the setting, and text type. An example of a context visual would be a photo that depicts a man and a woman talking to each other in a classroom. According to Ginther (2002), two main purposes of context visuals are (a) to set the scene for the verbal exchange, and (b) to indicate a change of speakers in a conversation.

Content visuals are visuals that are related to the content of the verbal interaction and may include still photos, pictures, drawings, diagrams, etc. A photo of Leonardo DaVinci's *Mona Lisa* accompanying a lecture on arts is an example of a content visual. Bejar et al. (2000) classify all content visuals into four groups: content visuals that replicate the audio stimulus, content visuals that illustrate the audio stimulus, content visuals that organize information in the audio stimulus, and content visuals that supplement the audio stimulus. All groups of content visuals can be represented either graphically or textually or both graphically and textually. According to Bejar et al. (2000), the first three types of content

visuals facilitate the comprehension of the oral stimulus, while the last type of content visuals makes it harder.

The effect of visuals (i.e., facilitating vs. debilitating) on learning may depend on various factors (Chung, 1994; Schriver, 1997). Chung (1994), for example, notes that the visual information becomes facilitative when the language learner can interpret its meaning correctly; otherwise, the visual information can be distracting for the learner. Schriver (1997) argues that overall pictures are beneficial for language learners because memory for pictures appears to be better than memory for words. However, she also acknowledges that pictures can be distracting in those cases, when they decorate the text and do not convey any meaningful information. Therefore, it seems very important "to bring words and pictures together in harmonious ways" (Schriver, 1997, p. 411).

Taking into account that L2 learner's comprehension can be affected by the correspondence between what is said and what is seen in a picture or video, Schriver (1997) proposes five ways, in which textual and visual information can be integrated.

1. Redundant – when words and pictures convey identical content.
2. Complementary – when words and pictures provide different content, with both modes being necessary to understand the main idea.
3. Supplementary – when words and pictures provide different content, with one mode presenting the main idea and the other mode supplementing it.
4. Juxtapositional – when words and pictures provide different content, with both modes presenting the ideas that clash; the main idea can be inferred only when both modes are presented simultaneously.

5. Stage-setting – when words and pictures present different content, with one mode providing the content and another mode giving the main idea.

Thus, context visuals and content visuals can have either facilitative or debilitating effect on learning depending on their interaction and correspondence with the accompanying textual or oral information.

Multimedia Language Learning

With an easy access to technology, it became possible to combine different types of media, such as texts, sounds, images, and video, which led to the development of multimedia learning. According to Chung (1994), multimedia is "the combination of two or more media such as text, images, or sound" (p. 1). Multimedia learning occurs when new information is presented to students in several modes, for example in pictures and words (Mayers, 1997).

A model that explains the interaction of textual and visual media by examining surface features of media, underlying cognitive functions, and characteristics of tasks and learners was proposed by Salomon (1989). According to this model, the process of learning from visual and textual media involves five types of variables (i.e. stimulus variables, cognitive variables, person variables, task variables, and accomplished psychological functions), integration of which he calls "visual supplantation" (Salomon, 1989, p. 77). Salomon's (1989) visual supplantation is a process when explicit visuals "model (that is – supplant) the kind of imagery that learners should have conjured up on their own, assuming… that such imagery is necessary for the acquisition of the material to be learned" (Salomon, 1989, p. 77). In other words, visual supplantation facilitates the process of learning only when there is a complementary relationship between the textual information and information presented in visual sources.

Salomon's (1989) idea of visual supplantation parallels Mayer's (1997) contiguity effect, according to which learners can build connections between verbal and visual stimuli more effectively when the text and illustrations are presented contiguously or simultaneously. Based on a series of studies, Mayer (1997) found that students with low background knowledge and high levels of spatial ability benefit most from the contiguous presentation of verbal and visual information. However, it should be noted that both Salomon's (1989) visual supplantation and Mayer's (1997) contiguity effect imply a relationship between the textual information and the information contained in *content* visuals, not *context* visuals. Taking this factor into account, Ginther (2002) hypothesizes that context-based visual stimuli can be distracting for learners when they look for content-based information.

Another theory that explicates possible interactions occurring among textual, visual, and individual variables in students' comprehension is Generative Theory of Multimedia Learning (Mayer, 1997). Mayer's (1997) generative theory of multimedia learning, which is based on Wittrock's (1974, 1989) generative theory and Paivio's (1986) dual coding theory, attempts to shed light on the learners' integration of visual and verbal information to comprehend a text. Mayer (1997) summarizes his theory in the following way:

> In a generative theory of multimedia learning, the learner is viewed as a knowledge constructor who actively selects and connects pieces of visual and verbal knowledge. The basic theme of a generative theory of multimedia learning is that the design of multimedia instruction affects the degree to which learners engage in the cognitive processes required for meaningful learning within the visual and verbal information processing systems. (p. 4)

Figure 2.2 depicts three main processes involved in multimedia learning: selecting words and images from the input, organizing them into visual and verbal mental representations, and integrating them. Mayer (1997) notes, however, that because of the limited capacity of short-term memory to hold large amounts of information, memory load may constrain the integration of visual and verbal information. Therefore, given these limitations of working memory, "the processes of selecting, organizing, and integrating are more likely to occur when visual and verbal information is presented contiguously rather than separately" (Mayer, 1997, p. 11).

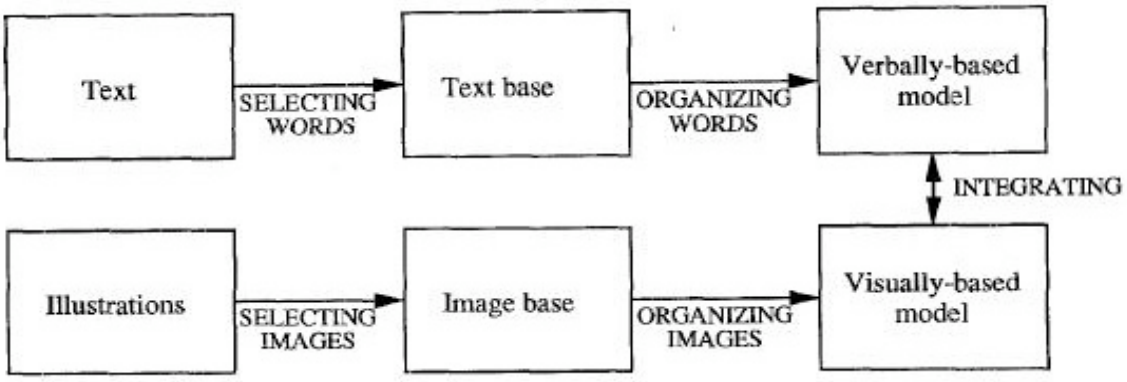

Figure 2.2. A generative model of multimedia learning (Mayer, 1997, p. 5).

Thus, the above-mentioned models of multimedia learning imply contiguity, simultaneity and integrity of textual and visual information for effective learning to occur.

Research on the Use of Visuals in L2 Listening Comprehension

Many L2 listening researchers appear to agree that L2 listening ability involves not only verbal input, but non-verbal components as well due to the fact that in the majority of the real-life listening situations the listener is able to see the speaker (Baltova, 1994; Gruba, 1997; Progosh, 1996; Wagner, 2007). Thus, they argue that the inclusion of the visuals in L2 listening tasks can assist language learners in processing and understanding verbal input and

can lead to increased performance. According to Wagner (2007), the use of the visual channel together with aural input can not only result in more authentic L2 listening tasks, but might also "lead to more construct relevant variance in the assessments, allowing for more valid inferences to be made from the results of those assessments" (p. 67).

Some of the advantages of visual input for the L2 language listener are as follows:

1. Seeing the situation and the participants increases situational and interactional authenticity (Buck, 2001; Wagner, 2007).
2. Body language, facial expressions, and gestures of the speaker can provide additional information (Buck, 2001; Coniam, 2001; Ockey, 2007; Rubin, 1995).
3. With visual input, the listener can more easily identify the role of the speaker and the context of the situation (Baltova, 1994; Gruba, 1997; Rubin, 1995).
4. Visual elements can activate the listener's background knowledge (Ockey, 2007; Rubin, 1995).

The effects of visuals on the listening comprehension depend on a number of different factors, such as the task, the types of visual materials used, the characteristics of the learners, and the interaction of these factors (Ginther, 2002).

Significant amount of research has been done on the role of video in L2 listening comprehension. The advantages of video over audio-only format are related to context, discourse, paralinguistic features, and cultural aspects (Coniam, 2001). These advantages appear to be more significant in a communicative approach to language learning. Secules, Herron, and Tomasello (1992) claim that video provides the learning experiences that are more likely to occur in the real world. They suggest that "video permits second language learners to witness the dynamics of interaction as they observe native speakers in authentic

settings speaking and using different accents, registers, and paralinguistic cues (e.g., posture, gestures)" (Secules, Herron, & Tomasello, 1992, p. 480).

According to Gruba (2006), "to date, no single definition of video-mediated listening comprehension has become established" (p. 77). Discussing the benefits of video, Gruba (2006) asserts that "visual media may foster macrostructure development by illustrating abstract concepts in a concrete way" and "assisting in the construction of mental models" (p. 79). Rubin (1995) argues that "appropriately selected video can be… the most facilitative environment for listening, especially at the elementary language learning stages" (p. 151).

A great deal of research on the use of visuals in listening comprehension emerged over the past years, but the results of these studies were often inconclusive. Some studies suggested that visual aids were more helpful for less proficient language learners, especially when the latter encountered difficult texts (Mueller, 1980; Rubin, 1995). In other studies, it was found that non-verbal clues have little, if any, effect on facilitating the understanding of verbal information (Lynch, 1998).

A number of research studies on the use of visuals that were conducted during the last decade involved second language learners of French (Baltova, 1994; Chung, 1994; Jones, 2002, 2003; Secules et al., 1992). In some of these studies, it was found that video-based materials were effective for the development of listening comprehension (Baltova, 1994; Chung, 1994; Secules et al, 1992) and "generated positive attitudes and confidence in understanding even in the case of poor comprehension" (Baltova, 1994, p. 520). The use of images also appeared to promote listening comprehension (Baltova, 1994; Chung, 1994), but the findings of Chung's (1994) study suggested that the use of multiple images could be distracting for language learners.

Multimedia learning and multimedia environment were also the focus of some research on L2 listening (Brett, 1997; Jones, 2002, 2003). Brett (1997), for example, found that computer-based multimedia environment resulted in more effective development of listening comprehension skills and better performance on language recall tasks. Mayer's (1997) generative theory of multimedia learning was the focus of Jones's (2002) study that investigated the influence of visual and verbal annotations on listening comprehension of second language students of French. While the students with access to both visual and verbal annotations performed best, the students with no annotations available showed the lowest results on understanding the passage and learning the vocabulary (Jones, 2002). The interviews with test-takers conducted by Jones (2003) corroborated the results of her 2002 study and provided "qualitative evidence for a generative theory of multimedia learning that suggests that the availability and the choice of visual and verbal annotations in listening comprehension activities enhances students' abilities to comprehend the material presented and to acquire vocabulary" (p. 41).

Overall, as the results of the existing studies show that visuals can be both facilitative and distracting for L2 listening comprehension, most of the L2 listening researchers call for more studies to better understand the role of visual information in listening comprehension.

Research on the Use of Visuals in L2 Listening Tests

Several studies investigating the use of visuals in listening tests have been carried out during the last two decades (e.g., Coniam, 2001; Ginther, 2002; Gruba, 1993; Jones, 2003; Ockey, 2007). However, there is a dearth of research on visual support in testing L2 listening, specifically studies comparing audio-only listening tests with tests that include images and video. According to Buck (2001), the most vital question that needs to be

considered is whether the use of visuals makes a difference to test-takers' listening comprehension. In addition, Buck (2001) proposes several important caveats for the developers of listening tests:

> …when testing language ability, the emphasis needs to be on processing linguistic information, not visual information. Furthermore, it seems sensible to bear in mind that adding visual information is probably only worthwhile if it provides us with better assessments of the listening construct. In some cases the visual information may serve to increase the cognitive load of the test-taker, and that may interfere with the testing process. (pp. 253-254)

Some researchers hypothesize that audio-based listening tests will eventually be replaced by computer-based listening tests that will include at least some type of visual aid (Ockey, 2007). One of the studies, conducted by Coniam (1999), examined test-takers' reactions to computer-based tests as opposed to pen-and-paper tests. Based on the results of the study, Coniam (1999) argued that test-takers need to be familiar both with the computer in general and with the test types before they can be expected to take computer-based tests. While computer-based tests with multiple-choice questions appeared to be acceptable by the participants of the study, Coniam (1999) found that test-takers responded less favorably to tasks, which required more than just a mouse click. Therefore, he suggested that the reliability of the results obtained from computer-based tests may be questionable when test-takers are given tasks on computer that differ from the tasks in a pen-and-paper mode.

Talking about the rationale for using visuals in computer-based listening tests, Ockey (2007) argues for the use of at least some sort of visual stimulus because "most target-language-use situations include visual stimuli, and it seems unreasonable to have test takers

stare at a blank computer screen while taking a computer-mediated test" (p. 517). The main advantage of video is its authenticity, while still images are easier and cheaper to produce and incorporate in the listening tests. However, the role of video as a listening assessment tool is less clear and research in this area has yielded inconclusive results (Coniam, 2001).

Discussing the role of video media in the assessment of listening skills, Gruba (1997) proposed four reasons for using video in listening assessment. First, the use of video is *theory driven* because, according to models of language comprehension, communication in real-life situations involves both verbal and visual elements. Second, the use of video in listening assessment is *pedagogy related* since language instructors always incorporate visual aids in their teaching for a number of pedagogical reasons. Third, video media should be used in language assessment because "there are features of the process, or setting, of how the language is being used which cannot be separated from its meaning" (Gruba, 1997, p. 339). Finally, the fourth reason for using video media in the listening assessment is justified by their existing use in distance learning programs.

Computer-based testing can effectively represent and test communicative competence targeting the test-takers' ability to use language in different contexts. However, the existing definitions of context seem to be too broad and all-inclusive, thus allowing multiple interpretations on the part of researchers (Lynch, 1998; Ginther, 2002). This multiplicity often results in researchers not knowing which aspects of context to focus on when creating and using context visuals for listening tests. In TOEFL CBT, for example, context visuals represent two features of context: situation and participants. However, the effects of the representation of context in specific testing situations remain undetermined (Ginther, 2002).

As mentioned by Ginther (2002), the TOEFL CBT Listening Comprehension section uses mostly context visuals. The following four combinations of audio and visual stimuli can be found in the Listening Comprehension section of TOEFL CBT: (a) dialogues accompanied by a context visual (for example, a picture of a speaker and setting); (b) short conversations with context visuals (for example, a series of pictures of the speakers and setting); (c) a combination of academic discussions and context visuals (for example, a series of pictures of two students and setting); and (d) mini-talks complemented by context visuals (for example, a series of pictures of two interlocutors) and/or content visuals (for example, a diagram related to the content of the conversation between two interlocutors).

Existing listening tests can employ five possible modes of input: audio-only, context-only still images, context-only video, content still images, and content video. Ockey (2007) suggests that depending on the mode of input used in listening tests, test-takers can process verbal information in different ways and, therefore, perform on tests differently. This, in turn, might affect the construct validity of listening tests.

The results of the existing comparative studies on the role of visuals in listening tests appear to be inconclusive. While some studies showed that visuals can improve students' performance on listening tests (e.g., Ginther, 2002), others evinced no facilitative effect of visuals on listening comprehension of test-takers (e.g., Coniam, 2001; Gruba, 1993; Ockey, 2007). This discrepancy of the results might probably be due to the use of different types of visuals, namely content and context visuals.

In his study, Gruba (1993) administered the same academic lecture in video and audio mode to 91 advanced-level ESL students and compared students' performance on the two modes of presentation. The results of the study did not find any statistically significant

difference between students' scores for a video-mediated test and an audio-mediated test. One of the possible problems that might have affected such results was low reliability of the test (.45). Additionally, Gruba (1993) hypothesized that no differences between the video and the audio groups could be due to the possibility that advanced second language learners are not "medium-dependent" (p. 87). Finally, Coniam (2001) suggested that as the video version of the test in Gruba's (1993) study was broadcast through a single video monitor, such mode of display could have influenced test validity because test takers had to constantly look up and down and shift their focus of attention between the screen and the question paper.

Interestingly, Gruba (1997) moves away from the comparative approach (i.e., when an audio-only test is compared to a video-mediated test) that he employed in his 1993 study and claims that it should be abandoned. Instead, he advocates Salomon's (1991) idea to use of a combination of "systemic and analytic" methods with "prolonged observations, a clear statement of goals and sophisticated research designs" (Gruba, 1997, p. 340).

Another important comparative study of two modes of a listening test (i.e., video and audio-only tests), which included test items in an open-ended format that required short or extended answers, was carried out by Coniam (2001). After completing the tests, the 104 test takers, who were Hong Kong English language teachers, filled out a questionnaire that included questions about advantages and disadvantages of video and audio and test-takers' preference for either mode. The analysis of test-takers' scores showed that the audio group performed better on the listening test than the video group; however, this difference was not statistically significant. Furthermore, the results of test-takers' responses to the questionnaire indicated that while the video group did not consider a video-mediated listening test to be of any advantage, the audio group did not perceive any advantage either. Coniam's (2001)

believes that the possible reason for the video-mediated listening test not being advantageous is due to the nature of the used videotext that was in the form of a talk show (i.e., context video). Thus, the results could have been different, had the video contained more than just paralinguistic clues. Finally, Coniam (2001) suggests:

> … it is likely that in different contexts, people may prefer different 'mediums'. For example, when taking a listening test, a participant may opt for the audio mode but the very same person might opt for the video mode while learning a language. (p. 12)

There appears to be a dearth of studies that compared different types of visuals (i.e., content vs. context visuals) and examined their effect on test-takers' performance (Ginther, 2002; Ockey, 2007).

Ockey (2007) compared context-only still images to context-only video in academic computer-based listening test and looked at the ways test-takers engaged with these two modes of input. Six ESL students categorized into three ability levels took two listening tests based on the same two-hour lecture: one listening test was video-mediated, the other test included a series of five still images. Moreover, interviews, retrospective verbal reports and videotaping of the participants were used to collect additional information. While all test-takers engaged minimally with still images, their engagement with the video stimulus was more extensive. The results indicated that context-only still images were helpful at the beginning of the listening test because they provided situational context, but were of no help later during the test. As for the video stimulus, while it was helpful for some students, it was found to be of no help and even distracting for others.

Additionally, Ockey (2007) analyzed the following five types of visual cues: lip movements (lip reading), gestures to indicate topic change, hand motions to signal key

words, facial gestures to indicate opinion, and body gestures to indicate emphasis. He discovered that the use of visual clues was helpful for some test-takers but distracting for others. It seems that two reasons might have caused this discrepancy: learning style preferences (e.g., unlike auditory learners, visual learners could benefit from videos a lot more), and cultural differences (e.g., video might be distracting for students from cultures, where looking into somebody's eyes is considered inappropriate).

There are three main implications for L2 test developers on the basis of Ockey's (2007) findings. First, test developers should use one still image rather than a series of images in listening tests because some test-takers seem to be distracted by the changes of images within one test. Second, still images appear to have the minimal bias against test-takers who do not use visual clues during listening tests. Third, the use of the video stimulus in listening tests may measure test-takers' ability to use visual clues to facilitate their listening comprehension.

The effects of the presence or absence of different types of visuals, type of stimuli (Dialogues/Short Conversations, Academic Discussions and Mini-talks), and language proficiency on students' performance on TOEFL CBT listening comprehension section were the emphasis of Ginther's (2002) study. The study involved 160 ESL students who were administered 40 listening comprehension items, 20 of which were accompanied by visuals and 20 were not. The results of the study revealed that facilitation occurred with content visuals accompanying Mini-talks and context visuals accompanying Academic Discussions. In particular, one of the main conclusions drawn by Ginther (2002) was that "facilitative effects occur when the presentation of visuals in contiguous and, most importantly, when it is directly related to the content of the information presented in the audio portion of the stimuli

or it marks a turn in the conversation" (p. 162), thus supporting Salomon's (1989) and Mayer's (1997) theoretical discussions of contiguity and complementation. Finally, she called for further research of content visuals, especially their putative potential to facilitate comprehension of longer and more difficult audio stimuli.

Investigation of different text types (i.e., academic lecture and dialogue) was also the focus of a study conducted by Wagner (2007), who compared the test-takers' orientation to the video monitor for academic lecture and dialogue. The participants were videotaped during the listening test to compute the amount of time they oriented to the video monitor. It was found that overall test-takers looked at the video monitor 69% of the time. In addition, dialogues were viewed more frequently (72% of the time) than lectures (67% of the time). Wagner (2007) explains these results by claiming that the context-embedded nature of dialogues makes them more useful for learners because of the prevalence and salience of contextual cues and non-verbal elements in an interaction involving two speakers. In contrast, lectures are more context-reduced, less interpersonal, and contain the verbal input from only one speaker, with non-verbal elements being less prevalent and salient.

There are also some studies that explored L2 students' preferences of visuals in listening tests and examined their viewing behavior (e.g., Progosh, 1996; Wagner, 2007). Progosh (1996), for example, administered a video-mediated quiz and a survey questionnaire to Japanese learners of English to determine their opinions of the use of video media in listening tests. The results of the survey revealed that 91.9% of the students preferred video-mediated listening quizzes to audio-only quizzes.

Wagner's (2007) study appears to be the only study to date that examined test-takers' viewing behavior on an L2 listening test, namely the amount of time participants made eye

contact with the video monitor while the aural input was being presented. He suggests that in order to determine whether video can be a distraction or not, it is first necessary to investigate the extent to which test-takers watch the video during the listening test. Similar to Wagner (2007), Wetzel, Radtke, and Stern (1994) concur that it is difficult to determine exactly what learners attend to when watching video. They state:

> Because learners are less restricted in the information they may extract from images, meaning in a visual image tends to be inherently more ambiguous than that of an "equivalent" verbal message. In the absence of auxiliary direction, learners may ignore important elements of the intended message, or may attach spurious significance to the extraneous parts of the image. (Wetzel, Radtke & Stern, 1994, p. 182)

To conclude, no agreement among researchers seems to exist on whether visuals make a difference for test-takers' performance on L2 listening tests as the existing studies of visuals in listening tests yield inconclusive results. Therefore, more research is needed to determine how different types of visuals and different text types affect listening test scores, as well as what role learning style differences, cultural differences, and background knowledge play for test-takers' results on L2 listening tests.

Issues with the Construct Definition of L2 Listening Ability

Many test developers avoid using video-mediated tests because of the possible problems with their construct validity (Progosh, 1996). The commonplace concern is whether such listening tests measure what they purport to measure, or whether they measure some other aspects that may affect test-takers' scores. Ockey (2007) urges that in order to use video in computer-based listening tests, test developers must clearly define the listening

construct they purport to measure. He argues that the listening construct needs to be enlarged so that it would include the ability to obtain information from visual clues and even the ability to take notes. Thus, a number of researchers claim that with the implementation of visuals in listening assessment, the validity, usability, and reliability of such listening tests must be investigated and rethought (Gruba, 1997; Ockey, 2007; Wagner, 2007).

The role of non-verbal information is usually ignored in construct definitions of L2 listening ability (Buck, 2001; Gruba, 1993; Wagner, 2007). Buck (2001), for example, argues that in L2 listening tests, where the listener is only the recipient of aural input and is not involved in interaction, it is better to avoid the use of video and keep the emphasis on testing language ability. Moreover, as different people utilize visual information differently, it is better to focus only on the audio information. Buck (2001) states:

> My own instinct is that visual information is more important in interactional language use, where the emphasis is on the relationship between the participants, but less important in transactional language use where the emphasis is on the content. In most cases, language tests are assessing transactional language use. Furthermore, we are usually interested in the test-takers' language ability, rather than the ability to understand subtle visual information. (p. 172)

According to Buck (2001), the only possible compromise in L2 listening testing is to present a still picture of the participants and the scene instead of the video.

However, taking into account that non-verbal information is an integral part of interpersonal communication in a real-life situation, other L2 researchers argue for the inclusion of non-verbal components in the construct definition of L2 listening ability in listening tests (e.g., Progosh, 1996; Wagner, 2007). Wagner (2007) suggests that listening

tests are designed to measure test-takers' listening skills "in a communicative language ability framework" (p. 68). As the verbal and non-verbal channels are "inextricably intertwined in the communication of the total meaning of an interpersonal exchange" (Burgoon, 1994, p. 347, cited in Wagner, 2007, p. 69), the exclusion of non-verbal information from listening tests might threaten their validity. Progosh (1996), also supporting the use of video in listening tests, claims that "paradoxically, the validity of listening tests that do not take into account that most people both hear and see in most communicative situations is just … contentious" (p. 35).

Thus, with the apparent lack of unanimity among researchers on the construct definition of L2 listening ability, it appears that further research is needed to ensure that listening tests, both with and without visuals, measure the right construct.

CHAPTER 3. METHODOLOGY

The purpose of Chapter 3 is to provide the information about the participants of this research project, materials that were developed and used in the study, the setting where the study took place, and procedures. Additionally, this chapter gives a description of the pilot study and the data collection for the main study. Finally, Chapter 3 concludes with a discussion of the data analysis employed to answer the research questions of the study.

Participants

The participants for the main study were non-native speakers of English enrolled in three listening classes at Iowa State University (ISU): one high-level listening class with students from Intensive English and Orientation Program (IEOP), which is a pre-university program, and two listening classes (99L) with students enrolled in regular classes at ISU. Overall, taking into account that IEOP class consisted of students who did not obtain the score on the TOEFL test high enough to be enrolled in regular undergraduate classes, while 99L classes included students who had passed the TOEFL test and were enrolled in regular classes at ISU, the overall English proficiency level of IEOP students was considered to be lower than that of 99L students. For the purposes of convenience in this study, IEOP students will be called IEOP group, and students from two 99L classes will be called 99L1 group and 99L2 group respectively.

To invite students for participation in the study, the researcher visited three listening classes, where he presented the purpose of the study to the students and answered their questions related to the procedures of data collection. The students willing to participate were given sufficient time to read and sign the consent forms.

The total of 34 students participated in the study, including 12 students from IEOP group, 13 students from 99L1 group, and 9 students from 99L2 group. Table 3.1 provides participants' profile with the background information about them.

Table 3.1

Participants' Profile

Student #	Group	Age	Gender	Native language	Years of learning English	Time in the US (in months)	Placement test listening scores[a], %
1	IEOP	45	f	Korean	10	12	65
2	IEOP	19	f	Chinese	8	2	75
3	IEOP	29	f	Tatar, Russian	3	9	85
4	IEOP	25	m	Japanese	3	2	35
5	IEOP	19	m	Chinese	10	2	80
6	IEOP	19	m	Chinese	6	2	85
7	IEOP	18	m	Chinese	6	2	85
8	IEOP	18	m	Chinese	6	2	85
9	IEOP	22	m	Arabic	17	2	95
10	IEOP	20	f	Korean	-	12	-
11	IEOP	19	m	Chinese	4	7	65
12	IEOP	26	m	Chinese	2	5	75
13	99L1	18	m	Chinese	10	4	63
14	99L1	18	m	Chinese	9	7	47
15	99L1	18	f	Laos	8	4	35
16	99L1	19	m	Chinese	4	2	27
17	99L1	20	m	Chinese	6	3	43
18	99L1	26	f	Korean	6	12	40
19	99L1	20	m	Chinese	6	2	43
20	99L1	20	m	Chinese	10	12	48
21	99L1	19	m	Chinese	8	7	63
22	99L1	19	m	Chinese	7	9	37
23	99L1	20	m	Chinese	10	24	37
24	99L1	19	m	Chinese	0.5	2	20
25	99L1	19	f	Chinese	6	2	33
26	99L2	19	m	Korean	10	8	43

Table 3.1 (continued)

Participants' Profile

27	99L2	18	m	Chinese	3	2	37
28	99L2	18	m	Chinese	8	3	20
29	99L2	20	m	French	2	3	47
30	99L2	22	m	Chinese	3	2	43
31	99L2	18	f	Chinese	3	2	47
32	99L2	19	f	Chinese	8	2	47
33	99L2	20	m	Chinese	5	2	-
34	99L2	20	m	Chinese	6	2	30

Note. [a]Listening scores for IEOP group are from Michigan Test Battery (MELAB), and listening scores for 99L1 and 99L2 groups are from English Placement Test (EPT).

From the information presented in Table 3.1, it can be seen that that the majority of the students were 18 to 20-year old native speakers of Chinese. Out of 34 participants, nine were females and 25 were males. Most of the participants were in the USA only for several months and only five students lived in the USA for a year or more.

Materials

For the purposes of this study, the researcher designed a computer-based Listening Test (LT) that consisted of six listening passages (LP) and 30 multiple-choice questions (five questions for each passage). The Listening Test (see the script in Appendix A and screen shots of an online version in Appendix B) started with the instructions page that test-takers could both read on the screen and listen to in an audio format. Overall, the duration of the Listening Test was about 44 minutes and the difficulty level of the listening passages was similar as determined by expert analysis discussed later in this chapter. Internal consistency reliability (KR-20) of the Listening Test was .70, which, considering the relatively small number of participants in this study, is acceptable.

As most of the existing studies of visuals in listening tests included short conversations/dialogues and/or academic lectures (e.g., Coniam, 2001; Ginther, 2002; Ockey, 2007; Wagner, 2007), in this study each listening passage utilized one of the following two text types: a dialogue between two college students or a professor and a student (D) or a short academic lecture given by a university professor (L). In addition, on the basis of suggestions from previous research on visual support in listening tests (e.g., Chung, 1994; Ockey, 2007), the researcher incorporated one of the three types of visual input - a single photograph, video, or no visuals (i.e. audio-only format) - in each listening passage of the Listening Test. Table 3.2 outlines the structure of the Listening Test.

Table 3.2

Structure of the Listening Test

Audio-only format (A)		Photograph (P)		Video (V)	
Dialogue 1	Lecture 1	Dialogue 2	Lecture 2	Dialogue 3	Lecture 3
LP1	LP4	LP6	LP2	LP3	LP5

Note. LP – listening passage.

To ensure randomness of the visual input and text type, the listening passages were administered in the following order: AD → PL → VD → AL → VL → PD (or LP1 → LP2 → LP3 → LP4 → LP5 → LP6).

According to Rubin (1995), selection of texts for listening passages should consider the amount of background knowledge (such as cultural, linguistic, and world knowledge) required as it can affect listening comprehension of test-takers. Therefore, the texts of the listening passages written for the Listening Test covered general topics in Journalism, Linguistics, Biology, Sport and Nutrition, and History that do not require prior specialized knowledge in those areas.

Research suggests that the optimal length of a listening passage for beginning and intermediate L2 learners is between thirty seconds and two minutes (Rubin, Quinn, & Enos, 1988; Thompson & Rubin, 1993, cited in Rubin, 1995). As the participants of this study were intermediate to advanced ESL learners (based on the evidence from their placement test scores), the length of the listening passages varied from 2.5 to 3.5 minutes. Moreover, when designing the listening passages for the Listening Test, the researcher had to take into account the fact that each listening passage would be followed by five multiple-choice questions. Therefore, it would have been difficult to create five questions, had a listening passage been too short.

As looking up and down from question paper to screen may be one of the possible reasons why test-takers become distracted by the visual images (Coniam, 2001), in the given study it was decided to exclude this factor by having test-takers answer test questions on a computer screen. Thus, each listening passage was followed by five multiple-choice questions that were displayed on a computer screen. Test-takers had 12 seconds between the questions to choose the correct answer.

Although there is empirical evidence that question preview can facilitate test-takers' performance (e.g., Chang & Read, 2006), in this study test-takers could only hear the questions and multiple choices and then choose the letter that corresponded to the best answer. As it was the test of L2 listening comprehension, the researcher decided not to display the content of questions and answers on the screen in a written form; otherwise, the reading skills might have come into play and could have affected test-takers' performance and, consequently, the validity of the Listening Test.

As can be seen from the Table of Specifications (see Appendix C), the following five types of questions were used in this Listening Test: true-false (e.g., According to the passage, which of the following is true about pyramids?), exception (e.g., According to the passage, all of the following describe cacti EXCEPT), inference (e.g., What can be inferred about protein?), details (e.g., What is the student's main problem?), and purpose (e.g., Why does the man want to talk with the woman?).

Two graduate ISU students who were native speakers of English were asked to participate in audio- and video-recording of the listening passages, as well as a photo session to make pictures for two listening passages. Audacity, free open source software for recording and editing sounds, was used to record two lectures (LP1 and LP3) and two dialogues (LP2 and LP4), as well as all 30 multiple-choice questions, in mp3 format. A high-quality digital video camera was used for video recording of a lecture (LP5) and a dialogue (LP6). Both audio and video files were edited using Windows Movie Maker.

As mentioned above, four out of six listening passages used visuals: two listening passages with a single photograph in each passage and two video-mediated listening passages. The two photographs taken for the listening passages included a photograph of a lecturer for the lecture and a photograph of two speakers for the dialogue. The two videos included a video of a lecturer reading a lecture and a video of two students talking in the hall. Thus, according to the classification of visuals proposed by Bejar et al. (2001) and Ginther (2002), only context visuals were used in the Listening Test.

Adobe Dreamweaver CS3 was used to create a computer-based version of the Listening Test that could be administered in an online mode. The Listening Test was hosted

on one of the public servers at Iowa State University. The Firefox web browser was used to run the Listening Test.

It was also decided to use screen capturing software, Camtasia Studio 4, to record the test-takers' responses to the multiple-choice questions. The use of Camtasia also allowed to analyze the test-takers' behavior on the screen during the Listening Test.

Besides the Listening Test, a Pre-Test Questionnaire, a Post-Test Questionnaire, and a Perceptual Learning Preferences Survey were created to collect data from the participants of the study. The Pre-Test Questionnaire, consisting of 14 questions, was designed and used to obtain information about participants' age, native language, country of origin, educational background, time of exposure to English, time spent in the USA, and other information relevant to the study (see Appendix D).

To obtain information about students' learning styles, an adapted version of Kinsella's Perceptual Learning Preferences Survey (a self-rated questionnaire on learning-styles preferences, Kinsella, 1995) was used (see Appendix E). The Perceptual Learning Preferences Survey consisted of 24 questions, the main purpose of which was to determine test-takers' preferences for three learning styles: visual-verbal, visual-non-verbal, and auditory. Unfortunately, like in Chung's (1994) study, it appeared to be impossible to adequately classify the participants by any particular learning style. Specifically, the original idea was to correlate visual-verbal, visual-non-verbal, and auditory learning styles with test-takers' performance on video-mediated, photograph-mediated, and audio-only listening passages respectively. However, as photograph-mediated listening passages did include verbal component (i.e. audio input), such correlation would have been inaccurate. Thus, the information about test-takers' learning styles was not used in the statistical analysis of data.

The Post-Test Questionnaire was designed and used to get feedback on the Listening Test from the participants (see Appendix F). It included 15 questions asking test-takers' opinions about the usefulness of visuals and their preferences of visuals in the Listening Test, their note-taking and screen-watching behavior, and the perceived difficulties of the Listening Test.

Setting

The study was carried out in one of the computer labs in the English Department at Iowa State University. The lab was equipped with 16 high-performance computers running Windows XP and connected to the Internet. All the computers were arranged along the walls with monitors facing the center of the lab. Every computer had a large monitor and headphones so that each test-taker could hear the listening passages and questions individually.

Procedures

Study Approval

Prior to the launch of the research project, the University IRB committee approved the study. According to the IRB policy, every participant of the study must carefully read and sign the consent form that provides information about the study, its purpose, risks, benefits, procedures, and measures of ensuring the confidentiality of participants' data and protection of their privacy.

Expert Analysis and Pilot Study

Expert analysis was conducted to check the reliability and validity of the Listening Test. Several professors in the Department of English at Iowa State University were asked to

analyze the Listening Test, its appropriateness for the proficiency level of the test-takers, as well as its reliability and validity.

A pilot study was conducted to check the effectiveness of procedures related to test administration, clarity of instructions and questions, quality of audio and video recording, appropriateness of listening passages, and time constraints of the study. Three international students whose overall profile was similar to the profile of the students in the main study participated in the pilot study. The main finding of the pilot study was that the Listening Test should not be administered online due to some technical issues pertaining to the speed of the Internet connection and the use of refresh tags in each page of the Listening Test. In particular, the pages that contained video files up to 50 MB each took up to 40-50 seconds to upload, depending on the speed of the Internet connection. Thus, it was decided to burn the Listening Test on DVDs and administer the test locally from DVDs rather than online.

Pre-Test Procedures

The researcher had to meet with participants three times: before, during, and after the study. During the first meeting (a pre-test meeting), the researcher introduced the study to students and invited them to participate. Those students who agreed to participate were given consent forms to read and to sign and were provided with a short training. They were also asked to fill out the Pre-Test Questionnaire and the Perceptual Learning Preferences Survey.

Data Collection

The second time the researcher met with participants in the computer lab to administer the Listening Test. As the computer lab used for the study held only 16 computers, it was decided to administer the Listening Test to each group of the participants separately, resulting in a three-part data collection process.

Each participant was placed in front of a computer where the Listening Test was open in a separate window. Before asking students to start the Listening Test, the researcher checked whether every participant had signed the consent form and filled out the Pre-Test Questionnaire and the Perceptual Learning Preferences Survey. After giving participants the signed copies of their consent forms and turning on Camtasia, the researcher instructed each test-taker to put on the headphones and start the Listening Test.

As listeners, unlike readers, do not have an option to review the information that has been presented to them (Thompson, 1995), the participants of this study were given paper for taking notes. Note-taking allowed test-takers to jot down main ideas or facts from the lectures and dialogues that they could later use for answering questions. Furthermore, as multiple-choice questions were presented only in an auditory format, some test-takers in this study used the opportunity to write down questions and/or multiple-choice answers as they were listening to them and use them when providing responses on the screen.

After participants finished the Listening Test, the researcher collected their notes and asked them to fill out Post-Test Questionnaire. The responses of each test-taker and his or her behavior on the screen were saved as Camtasia files to be used for later data analysis.

Post-Test Procedures

The third time the researcher met with the participants of the study was during one of their Listening classes. The purpose of this last meeting was to give students their scores for the Listening Test and to provide them with feedback on their performance and their learning style preferences. Students were also given some treats and thanked for their participation in the study.

Analysis

Quantitative data were collected and used to answer the three research questions in this study. Test-takers' scores for the Listening Test were analyzed to answer the first two research questions. The results of Post-Test Questionnaires were used to address the third research question. To analyze test-takers' performance on the Listening Test, 34 Camtasia files containing test-takers' responses were viewed and the data were recorded and saved (see Appendix G). SAS statistics software was used to carry out the statistical analysis of the data.

Research Question #1

The first research question was addressed through the statistical analysis of test-takers' scores for the Listening Test. Specifically, descriptive statistics were calculated and an F-test from the ANOVA procedure followed by the Tukey-Kramer method for post hoc comparison was run to determine the difference in test-takers' performance on different types of visual input (i.e., a single photograph, video, and audio-only format). As each type of visual in the Listening Test was represented in different text types (i.e., a dialogue and an academic lecture), a one-sample t-test was used to determine whether test-takers performed differently on each text type. Due to the evidence for statistically significant difference between text types, the ANOVA procedure with 3x2 factorial structure (types of visual input by text types) was used to determine the variance in performance on different types of visuals for 34 test-takers. As with multiple comparisons the probability of making a type I error increases, the Tukey-Kramer post-hoc test was employed to adjust the probability value.

Research Question #2

To answer the second research questions, the researcher used the results of the F-test from the ANOVA procedure for types of visual input and the results of the ANOVA

procedure with 3x2 factorial structure (types of visual input by text types) followed by the Tukey-Kramer method for post hoc comparison.

Research Question #3

Test-takers' scores for the Listening Test and their responses from the Post-Test Questionnaires (see Appendix H) were used to answer the last research question. An F-test from repeated measures ANOVA procedure followed by a Tukey-Kramer method for post-hoc adjustment was employed to determine whether there was a difference between test-takers' performance on different types of visual input and their preferences of visual stimuli in L2 listening tests.

Table 3.3 summarizes all the data gathering methods used and types of data obtained in the study.

Table 3.3

Data Gathering Methods

Variables	Method	Data obtained
Independent	Placement test scores	Data on test-takers' listening proficiency level
	Pre-Test Questionnaire	Data on participants' background, age, L1, time of exposure to English
	Learning Preferences Survey	Data about participants' learning styles
Dependent	Listening Test scores	Data on participants' performance on the Listening Test
	Camtasia recordings	Data on test-takers' responses and behavior on the screen during the Listening Test
	Post-Test Questionnaire	Data on test-takers' perceptions about the usefulness of visuals in the Listening Test

CHAPTER 4. RESULTS AND DISCUSSION

Chapter 4 presents the results of data analysis used to answer the three research questions. This chapter includes a discussion of the findings regarding the influence of visual support on students' performance on the Listening Test, and participants' perceptions concerning the usefulness of visuals.

Research Question #1

Research Question #1 addresses the difference among three types of visual input concerning their impact on test-takers' scores. Table 4.1 presents descriptive statistics for each type of visual input (i.e. audio-only, a single photograph, and video) for 34 test-takers.

Table 4.1

Descriptive Statistics for Types of Visual Input

Type of visual input	Mean, $\bar{x}$	SD
Audio-only	6.35	1.98
Photograph	6.32	2.25
Video	5.06	1.74

Note. n=34.

Results shown in Table 4.1 reveal the highest mean for audio-only listening passages ($\bar{x}_A$=6.35). The mean for listening passages with photographs is slightly lower ($\bar{x}_P$=6.32), while the mean for video-mediated listening passages is the lowest ($\bar{x}_V$=5.06).

An overall F-test from the ANOVA procedure with visual input types as treatments was used to evaluate and compare treatment means for 34 test-takers. Assumptions about the validity of the ANOVA procedure (i.e., normal distribution of scores and equal variances)

were satisfied. The Tukey-Kramer method for post-hoc comparison was applied to adjust the p-value. The results of the F-test for the visual input effect are given in Table 4.2.

Table 4.2

F-test for Types of Visual Input

Visual input	Visual input	Differences of means	t-value	Pr > |t|	Adjusted p-value
A	P	0.03	0.07	0.9425	0.9971
A	V	1.29	3.19	0.0022	0.0061*
P	V	1.26	3.12	0.0027	0.0076*

Note. n=34. A – audio-only format, P – photograph, V – video. An asterisk * indicates statistically significant results at the p<.05 level.

As shown in Table 4.2, the difference between $\bar{x}_A$ and $\bar{x}_P$ is very insignificant. However, there is evidence that the differences between $\bar{x}_A$ and $\bar{x}_V$, and $\bar{x}_P$ and $\bar{x}_V$ are statistically significant at the adjusted p >.0061 and p >.0076 levels respectively. Thus, the means for audio-only listening passages and listening passages with a photograph are significantly higher than the mean for video-mediated listening passages.

As each type of visual input was represented in different text types, test-takers' scores on dialogues and academic lectures were analyzed using descriptive statistics (see Table 4.3).

Table 4.3

Descriptive Statistics for Text Types

Text type	Mean, $\bar{x}$	SD
Dialogue	10.12	2.47
Lecture	7.62	2.65

Note. n=34.

The results in Table 4.3 indicate that the mean for dialogues ($\bar{x}_D$=10.12) is higher than the mean for academic lectures ($\bar{x}_L$=7.62).

A one-sample t-test was run to determine if the mean difference in test-takers' performance on dialogues and lectures was statistically significant. The results of the t-test revealed that test-takers' performance on dialogues was significantly better than on lectures at the $p<.0001$ level.

As the difference between participants' scores for dialogues and lectures was statistically significant, this factor was taken into consideration when determining the variance in test-takers' performance on different types of visuals. Descriptive statistics for every listening passage are given in Table 4.4.

Table 4.4

Descriptive Statistics for Types of Visual Input by Text Types

Type of visual input	Text type	Mean, $\bar{x}$	SD
A	D	3.50	1.24
A	L	2.85	1.28
P	D	3.41	1.42
P	L	2.91	1.22
V	D	3.21	0.95
V	L	1.85	1.21

Note. n=34. A – audio-only format, P – photograph, V – video, D – dialogue, L – lecture.

As indicated in Table 4.4, the means for dialogues with all three types of visual input are higher than the means for lectures with three different visuals.

Table 4.5 presents the results of the ANOVA procedure with 3x2 factorial structure (types of visual input by text types) followed by the Tukey-Kramer method for post hoc adjustment for 34 participants.

Table 4.5

Results of the ANOVA for Types of Visual Input by Text Types

Visual input type by text type	Visual input type by text type	t-value	Pr > |t|	Adjusted p-value
AD	AL	2.48	0.0143	0.1371
AD	PD	0.34	0.7360	0.9994
AD	PL	2.25	0.0257	0.2203
AD	VD	1.13	0.2619	0.8701
AD	VL	6.30	<.0001	<.0001*
AL	PD	-2.14	0.0339	0.2726
AL	PL	-0.23	0.8221	0.9999
AL	VD	-1.35	0.1786	0.7560
AL	VL	3.83	0.0002	0.0025*
PD	PL	1.91	0.0574	0.3974
PD	VD	0.79	0.4318	0.9692
PD	VL	5.97	<.0001	<.0001*
PL	VD	-1.13	0.2619	0.8701
PL	VL	4.05	<.0001	0.0011*
VD	VL	5.18	<.0001	<.0001*

Note. n=34. A – audio-only format, P – photograph, V – video, D – dialogue, L – lecture. An asterisk * indicates statistically significant results at the p<.05 level. A positive t-value indicates that the mean for the first element of a pair is higher than the mean for the second element, and a negative t-value shows that the mean for the first element is lower than the mean for the second element of a pair.

According to the results of the ANOVA presented in Table 4.5, statistically significant difference (at the $p<.05$ level) between the means was found in the following five pairs: audio-only dialogue vs. video lecture ($p<.0001$), audio-only lecture vs. video lecture ($p<.0025$), dialogue with a photograph vs. video lecture ($p<.0001$), lecture with a photograph vs. video lecture ($p<.0011$), and video dialogue vs. video lecture ($p<.0001$). On the other hand, very little difference was observed between the following pairs: audio-only dialogue vs. dialogue with a photograph ($p<.9994$), audio-only dialogue vs. video dialogue ($p<.8701$), audio-only lecture vs. lecture with a photograph ($p<.9999$), dialogue with a photograph vs. video-mediated dialogue ($p<.9692$), and lecture with a photograph vs. video-mediated dialogue ($p<.8701$).

Thus, returning to the first research question (Is there a difference among different types of visual input, namely a single photograph, video, and audio-only format, in an L2 listening test in terms of their effect on L2 test-takers' performance?), the results of the data analysis suggest that, yes, there is statistically significant difference in test-takers' scores for listening passages with different types of visuals.

Research Question #2

Research Question #2 addresses the helpfulness of visual aids (i.e., a single photograph and video) for test-takers' performance on the Listening Test. To answer the second research question, the results of the ANOVA procedures reported in Tables 4.2 and 4.5 were used.

The results of the F-test in Table 4.2 indicate that test-takers' scores for video-mediated listening passages are significantly lower than the scores for audio-only passages ($p<.0061$) and listening passages accompanied by a photograph ($p<.0076$). In addition, as

shown in Table 4.5, for all the pairs with statistically significant difference at the $p<.05$ level between the elements, test-takers' scores for a video-mediated lecture were significantly lower than for listening passages with audio-only format and a single photograph. This evidence suggests that video was not facilitative for test-takers on listening tests.

However, it should be noted that the differences between the video-mediated dialogue and other four listening passages with audio-only format and a photograph were not statistically significant (with p-values varying from $p<.7560$ to $p<.9692$) and, in fact, judging from the t-values from Table 4.5, students performed slightly better on the video-mediated dialogue than on the audio-only lecture and the lecture with a photograph. This finding suggests that the use of video in dialogues does not have a detrimental effect on the scores of test-takers but, on the contrary, might even facilitate their performance.

As far as a photograph is concerned, the results of the F-test from Table 4.2 suggest that compared to audio-only listening passages the use of this visual in listening passages does not affect students' performance ($p<.9971$), but in comparison with video, the use of a photograph in a listening passage appears to be more effective at the $p<.0076$ level. However, as shown in Table 4.5, no significant difference was found between test-takers' scores for the dialogue with a photograph and the video-mediated dialogue ($p<.9692$) and the lecture with a photograph and the video-mediated dialogue ($p<.8701$). This finding suggests that the use of a photograph and video in a dialogue as compared to audio-only format does not make any difference in test-takers' performance.

Thus, returning to the second research question (Does the use of visuals, namely a single photograph and video, in an L2 listening test facilitate test-takers' performance?), the results of the data analysis suggest that the effect of visuals on students' performance

depends on a text type. In dialogues, the use of photographs and video does not affect test-takers' performance on the listening tests. In lectures, the use of photographs does not seem to make any difference in students' scores for the listening tests, but the use of video appears to be detrimental.

It should be noted, however, that according to the classification of visuals proposed by Bejar et al. (2000) and Ginther (2002), both photographs and video used in the Listening Test for this study were context visuals. As mentioned in Chapter 2, the effect of visuals may depend on various factors. In case of context video that appeared to have a debilitating effect on test-takers' performance on a video-mediated lecture, several possible reasons might have led to such results. First, context video could have been distracting for test-takers if they were looking for content-based information (Ginther, 2002). Second, unlike dialogues, lectures have reduced context, are less interpersonal, and contain the verbal input from only one person, with few contextual clues. Taking this factor into account, context video used in a lecture might not have any facilitative effect but, on the contrary, could even be a distraction for listeners. Finally, as no item analysis was conducted for the Listening Test, it is possible that questions for the video-mediated lecture were overall more difficult than the rest of the multiple-choice questions on the Listening Test.

Research Question #3

The last research question addresses test-takers' preferences of visuals for listening comprehension and the correspondence of these preferences to test-takers' actual performance on different types of visual stimuli.

The analysis of test-takers' responses to the Post-Test Questionnaire indicated that most test-takers preferred audio-only format. Nine participants claimed that video aided their

listening comprehension and only five participants found photographs to be useful, including two students who considered both video and photographs to be helpful for their listening comprehension. With regards to preferences, almost half of all test-takers (i.e. 15 out of 34) claimed they would prefer a listening test without any visuals, 12 test-takers said they would prefer a video-mediated listening test, and seven test-takers would go for a listening test accompanied by a photograph. Thus, even though 56 percent (or 19 out of 34 test-takers) claimed they would prefer at least some type of visual support, only 35 percent (or 12 out of 34 test-takers) found visuals helpful on the Listening Test.

Table 4.6 presents the results of an F-test from repeated measures ANOVA procedure followed by a Tukey-Kramer post hoc method for 34 test-takers.

Table 4.6

F-test for Test-Takers' Preferences by Performance

Performance by preference	Performance by preference	t-value	Pr > |t|	Adjusted P
AA	AP	1.25	0.2149	0.9414
AA	AV	1.00	0.3182	0.9841
AA	PA	1.10	0.2738	0.9715
AA	PP	1.10	0.2767	0.9728
AA	PV	0.15	0.8814	1.0000
AA	VA	3.86	0.0003	0.0077*
AA	VP	1.87	0.0652	0.6357
AA	VV	1.54	0.1279	0.8325
AP	AV	-0.39	0.7009	1.0000
AP	PA	-0.53	0.5996	0.9998
AP	PP	-0.16	0.8721	1.0000
AP	PV	-1.08	0.2827	0.9748
AP	VA	1.28	0.2038	0.9330
AP	VP	0.65	0.5203	0.9992
AP	VV	0.05	0.9605	1.0000
AV	PA	-0.15	0.8814	1.0000
AV	PP	0.24	0.8139	1.0000
AV	PV	-0.99	0.3272	0.9857
AV	VA	1.99	0.0504	0.5572

Table 4.6 (continued)

F-test for Test-Takers' Preferences by Performance

AV	VP	0.98	0.3290	0.9862
AV	VV	0.62	0.5393	0.9994
PA	PP	0.37	0.7108	1.0000
PA	PV	-0.71	0.4827	0.9985
PA	VA	2.76	0.0076	0.1486
PA	VP	1.15	0.2548	0.9642
PA	VV	0.68	0.4960	0.9988
PP	PV	-0.93	0.3539	0.9901
PP	VA	1.44	0.1549	0.8790
PP	VP	0.81	0.4221	0.9962
PP	VV	0.20	0.8428	1.0000
PV	VA	2.84	0.0057	0.1237
PV	VP	1.68	0.0973	0.7568
PV	VV	1.60	0.1136	0.7985
VA	VP	-0.66	0.5103	0.9991
VA	VV	-1.45	0.1501	0.8719
VP	VV	-0.55	0.5859	0.9998

Note. n=34. A – audio-only format, P – photograph, V – video, D – dialogue, L – lecture. An asterisk * in the last column indicates statistically significant results at the $p<.05$ level. A positive t-value indicates that the mean for the first pair is higher than the mean for the second pair, and a negative t-value shows that the mean for the first pair is lower than the mean for the second pair.

Overall, as can be seen from adjusted probability values in Table 4.6, test-takers' preferences of visuals did not coincide with their performance on different types of visual aids in the Listening Test. The only statistically significant difference at the $p<.0077$ level was found between the scores for audio-only passages and the scores for video-only passages of test-takers who preferred audio-only format of listening tests. This finding indicates that the test-takers who preferred audio-only listening tests performed significantly better on

audio-only listening passages than on video-mediated listening passages. One of the possible explanations is that the test-takers who preferred audio-only format of listening passages belonged to an auditory type of learners, who could get easily distracted by visuals in the listening tests and, therefore, performed significantly better on audio-only listening passages.

Thus, the answer to Research Question #3 (Do test-takers' preferences of visual stimuli in listening tests correspond to their actual performance on different types of visuals in the Listening Test?) is no, test-takers' preferences of visuals do not correspond to their performance on these visuals. The only exception was test-takers with a preference for an audio-only format of listening tests, who performed significantly better on audio-only listening passages than on video-mediated ones.

CHAPTER 5. CONCLUSION

This last chapter presents the implications that can be drawn from the results of the given study, describes its limitations, and provides ideas for future research on visuals in L2 listening tests.

Implications

The study presented in this thesis investigated the effect of visual aids, namely a single photograph and video, on L2 test-takers' performance on listening tests. It also examined the correspondence of participants' preferences of visuals to their performance on listening tests with different types of visual support. The obtained results allow for several implications to be made.

First, the use of visuals in L2 listening tests does affect test-takers' scores. However, the magnitude of the impact of visuals on students' performance seems to depend on the types of visuals used in listening tests. While the use of a single photograph in a listening test as compared to a listening test without any visual support does not make any significant difference in test-takers' scores, the use of video stimulus appears to have a negative impact on students' performance. Hence, when deciding to use video in listening tests, test developers and L2 listening instructors should be cautious about the possible detrimental effect of video on students' performance on L2 listening tests.

Second, it appears that different text types have an effect on test-takers' results on listening tests as well. In dialogues, the use of photographs and video does not affect test-takers' performance on the listening tests. In lectures, the use of photographs does not seem to make any difference in students' scores for the listening tests, but the use of video appears

to be detrimental. Thus, more research is needed to determine whether the use of video in different text types affects students' performance on L2 listening tests differently.

Third, as students' perceptions concerning the helpfulness of different types of visuals do not coincide with their performance on listening tests accompanied by various types of visual support, it is important for test developers to take this factor into consideration when developing new L2 listening tests. To decide whether to use visuals in a listening test, it appears to be imperative not to rely on test-takers' subjective perceptions, but to base one's decisions on the results of objective research.

Another important implication concerns the use of computers for designing and delivering listening tests. Computer-based language testing has a great potential for the use of visual support, multimedia, and interactivity that be implemented in L2 listening tests. Moreover, in case of multiple-choice questions, it allows for immediate feedback on test-takers' performance.

Finally, it is important to emphasize that all the visuals used in the Listening Test for this study were *context* visuals that provided only information about the scene of the verbal interaction and, in case of a video-mediated dialogue, indicated a change of speakers in a conversation. As suggested by some researchers (Chung, 1994; Ockey, 2007; Schriver, 1997), the use of visual stimulus can have no facilitative effect or be even distracting for test-takers in those cases, when visuals do not convey any meaningful information or when there is no correspondence between what is said and what is depicted in a photograph or video. Hence, a comparative research study on the use of *content* visuals that illustrate or supplement the content of the verbal input in L2 listening tests might yield different results.

Limitations

The study described in this thesis has a number of limitations that are related to both the design of the Listening Test and the design of the study itself.

Some researchers argue that multiple-choice tests are less authentic than short-answer tests or tests requiring an extended answer (Hearst, 2000). Unlike short-answer tests, multiple-choice tests "lend themselves to test-taking strategies, which do not evaluate the student's understanding of the question" (Hearst, 2000, p. 31). Therefore, as the Listening Test designed for this study consisted only of multiple-choice questions, the results of the study could have been different, had the listening tasks been more complicated.

Authenticity of the Listening Test is another concern of this study. The audio texts used in this study were designed and produced by the researcher and thus might lack authenticity. As mentioned by Gruba (1997), authentic texts are generally more preferable in listening tests.

Another limitation of this study that might have affected the study results is related to the design of the computer-based test. As all five multiple-choice questions following each listening passage were displayed on a single web page, test-takers had an opportunity to go back and change their answers to previous questions while listening to the next question. In fact, the analysis of 34 Camtasia files with recordings of participants' behavior on the screen during the Listening Test revealed that 17 out of 34 students changed at least one answer to the previous question, including eight students who changed more than one answer. There were two cases when students skipped a question and went back to answer it after answering the rest of the questions. Moreover, in two other cases students accidentally missed a question (e.g., by answering Question 4 instead of Question 3) and later, after realizing the

mistake, went back to answer the missed question. Finally, one of the students did not seem to take the Listening Test seriously, randomly choosing the answers for several questions before hearing the four options. To avoid this limitation, a computer-based listening test must present each test-taker with only one question per page and allow moving to the next question only after the test-taker submits an answer to the previous question.

Another caution in interpreting the results of the study is related to the fact that no item analysis was carried out for 30 multiple-choice questions used in the Listening Test. As mentioned in Chapter 4, it is possible that overall low results on the video-mediated lecture can be due to the significantly greater difficulty of the questions for this listening passage.

Finally, the last limitation deals with an issue of homogeneity of the participants. As the participants in this study came from two different listening classes (namely, IEOP class and two 99L classes), their proficiency levels and listening skills in particular could be different as well, thus resulting in two heterogeneous groups.

Ideas for Future Research

On the basis of the results obtained in this study and implications, several ideas for future research arise. Firstly, as only context visuals were used in this study, further research is needed to examine the effect of content visuals on test-takers' performance on L2 listening tests and compare the two effects.

Secondly, as the findings of this study suggest, text types can have an impact on listening test results. Therefore, it is important to further investigate the interaction of different types of visuals (i.e. content and context photographs and video) with different text types, such as dialogues and academic lectures.

Thirdly, as indicated by previous studies (e.g., Chung, 1994; Ockey, 2007), students can perform differently on listening tests with a single picture vs. listening tests with multiple pictures. Designing a study that would take this factor into consideration would shed more light on this issue.

Fourthly, more research must be completed to detect the reasons for individual differences, such as the role of cognitive load within the test-takers' visual and acoustic information processing systems (Mayer, 1997), learning styles of test-takers, and their L2 proficiency levels.

Finally, the effect of visual support in L2 listening tests may depend on the types of test items. In addition to multiple-choice questions used in this study, another option would be to use open-ended questions. Additionally, an interesting approach would be to use "media inclusion," i.e. the use of graphics, video, and audio within an item or set of items in an L2 listening tests (Zenisky & Sireci, 2002, p. 348). Such multimedia can be employed for better illustration of a particular context, visualization of a problem, or evaluation of a specified construct. The findings of a research study that would use visuals not only in listening passages but also in test items would greatly contribute to the understanding of the effects visuals play in listening comprehension and testing.

APPENDIX A. THE LISTENING TEST

Instructions to the Listening Test

The purpose of this listening test is to test your ability to understand spoken English in academic settings. The listening test consists of six listening passages: three conversations between two people and three lectures. Each listening passage lasts for about three minutes. After listening to each passage, you will be given five multiple-choice questions. You may take notes while listening and use them to answer the questions. After each question, you will hear four possible answers. On your computer screen, choose the best answer for each multiple-choice question. You will have 12 seconds to answer each question. You will hear the listening passages and questions only one time, so please be very attentive while listening!

Now you will start the test!

Listening Passage 1

Listen to a conversation on campus between a professor and a student.

Man: Good morning, Professor Smith! Do you have a minute? I wanted to talk to you about my final project for our Biology class.

Woman: Sure. Come on in, Chris. So what did you want to talk about?

Man: You see, I've got a problem with the final project. I know that it should be a small research project, but I can't decide on a topic.

Woman: Well, think about something that would be of interest for you personally.

Man: Umm, all right… As a matter of fact, I really like cacti but… I am not quite sure what exactly I can research about them.

Woman: I see… Well, I suggest that you zero in on those aspects that make cacti unique, something that distinguishes them from other plants. Is there anything you find fascinating about these plants and yet difficult to understand?

Man: Um… I guess what makes cacti really unique is the fact that they can live in the deserts for months without any water supply and yet survive. And they are believed to have the most beautiful flowers in the world although I heard some of these plants bloom only once a century!

Woman: Well, here you go! Do research on what makes cacti such great survivors in arid regions and investigate their adaptability. You can also try to find information about their flowers and blooming periods.

Man: Yeah, that sounds like a great idea! Oh, thank you so much for your advice, Professor Smith! I really appreciate it!

Woman: You are welcome. Is there anything else I can help you with?

Man: Well, maybe... I was wondering whether you could allow me to retake that Bio test that we had last week. You see, I was really frustrated because I had spent so much time studying for it and I didn't do well on it because I was late for the class and didn't have enough time to answer all the questions.

Woman: Hm, why did you come late? What happened?

Man: You know, usually I take a bus to campus but that morning I missed my bus so I had to drive my car. Unfortunately, I ran out of gas on my way to campus so I had to pull over, leave my car, and walk all the way to school.

Woman: I see. Well, things happen, and knowing you and that you are a pretty dependable student, I think I can allow you to retake the test. Do you want to do that today?

Man: Oh yeah, thank you, that would be great!

Woman: Good, then how about you stop by my office at 3:45 p.m.? I have a meeting at 2 p.m. but hopefully it will have ended by 3:30 p.m.

Man: All right, I will be in your office at 3:45 p.m. And thanks again, Professor Smith!

Woman: See you then, Chris.

Now you will hear 5 questions. After you listen to each question, choose the best answer. You will have 12 seconds to answer each question.

Question 1: What is the main purpose of Chris's visit to Professor Smith?

A. To check whether Professor Smith is busy or not.

B. To ask Professor Smith for help with the final project for Biology class.

C. To talk with Professor Smith about unique plants.

D. To discuss what they are going to do during a meeting in the afternoon.

Question 2: According to the passage, which of the following is true about Chris?

A. He always misses his bus to campus.

B. He is a lazy student.

C. He is interested in cacti.

D. He does not know how to do research.

Question 3: What will most likely be the best title for Chris's final project for his Biology class?

A. Survival skills of cacti in a desert.

B. The most unique plants on the Earth.

C. Blooming periods of flowers.

D. Adaptability of beautiful plants.

Question 4: What does Chris ask Professor Smith's permission for?

A. Working on his research project.

B. Retaking his Bio test.

C. Attending a meeting.

D. Missing a class.

Question 5: According to the passage, each of the following is true about Professor Smith EXCEPT

A. Professor Smith considers Chris to be a dependable student.

B. Professor Smith will have a meeting in the afternoon.

C. Professor Smith allowed Chris to retake the Bio test.

D. Professor Smith didn't believe Chris's excuse for being late for the Bio test.

This is the end of Listening Passage 1.

Listening Passage 2

Listen to the beginning of a lecture in a history class.

Well, we have a lot to cover today so let's get going. As I told you last time, we will continue the discussion of the Second World War and specifically today we'll be talking about the German-Soviet war.

The scope of World War 2 increased greatly after Germany led by Adolf Hitler attacked the Soviet Union on June 22, 1941. The German-Soviet War, also referred to as the Great Patriotic War, lasted for almost four years and ended on May 9, 1945. It was the largest theatre of war in human history in terms of numbers of soldiers, military equipment, and

casualties. Actually, this military conflict was the deadliest one in history with over 30 million people dead as a result. Finally, this war was famous for its atrocities such as the Holocaust. Has everyone heard about the Holocaust? Well, the Holocaust is usually referred to the deliberate extermination of the Jewish population in Europe by the Nazi regime. As a result, nearly six million European Jews were killed during the Second World War.

OK, now back to the war itself. So, as I said, the German-Soviet war started with the invasion of the Soviet Union by Nazi Germany. Being the largest invasion in history that included over 3 million German soldiers, it had a codename Operation Barbarossa. The goal of Operation Barbarossa was to attack the Soviet Union by surprise and to conquer it in eight weeks. Despite initial success of the German armies, this operation failed. An eight-week war, as planned by Hitler, turned into a four-year war that ended with the defeat of Nazi Germany.

So why did Operation Barbarossa fail? There are many different reasons, but the main ones are the following. First, Adolf Hitler and the German High Command underestimated the potential of the Soviet Union. In particular, Germany underestimated the effective control of the Soviet government, the technical capacity of the Soviet Union, and the potential of the Soviet Union to mobilize vast military forces within a short time. Second, the Germans experienced problems with logistical planning. The German armies didn't have enough fuel supplies for a long combat, nor did they expect that the Russian road network was in such poor condition and would slow down the movement of the German armies. The last main factor that resulted in the failure of Operation Barbarossa was severe winter weather. As Germany planned to finish the war quickly, the German armies were not prepared for harsh winter conditions. Unlike Russians, Germans were short of winter clothes and fuel. In

addition, some of the German weapons didn't function properly when the winter temperatures dropped down to -30 F (that is -35 C). Because of a severe Russian winter, hundreds of thousands of German soldiers died of cold.

Now you will hear 5 questions. After you listen to each question, choose the best answer. You will have 12 seconds to answer each question.

Question 1: What can be inferred about the German-Soviet war?

A. It was part of the Second World War.

B. It lasted only eight weeks.

C. It was very successful for Germany.

D. It started earlier than expected.

Question 2: According to the lecture, which of the following characteristics of the German-Soviet war is NOT true?

A. It was the largest war in human history.

B. It was the deadliest war in human history.

C. It was the longest war in human history.

D. It was famous for its brutality.

Question 3: What does the Holocaust usually refer to?

A. The deaths of millions of people.

B. Extermination of the Jewish population in Europe.

C. Atrocities of the Soviet army.

D. Number of soldiers who died during the war.

Question 4: According to the lecture, all of the following factors resulted in the failure of Operation Barbarossa EXCEPT

A. Harsh Russian winter.

B. Bad Russian roads.

C. Lack of food.

D. Lack of fuel supplies.

Question 5: Why were the German armies not prepared for fighting Russians in winter?

A. They did not think Russians could attack them.

B. They did not want to fight in winter.

C. They did not know Russian roads well.

D. They did not count on such a long war.

This is the end of Listening Passage 2.

Listening Passage 3

Listen to a conversation on campus between two students.

Man: Hey, Mary, how's it going? I haven't seen you for a while.

Woman: Hi, Mark, it's going pretty well. I've been busy with my thesis.

Man: Oh… Are you graduating soon?

Woman: Hopefully. I was planning to do that next spring, but most likely it will be in summer. You know, this thesis project is killing me and it looks like I won't be able to finish and defend it by the end of the spring semester.

Man: What's the matter?

Woman: Well, in my thesis I am analyzing the extent to which protein influences the growth of muscles in humans.

Man: Hm, that sounds quite interesting...

Woman: Yeah, but I still can't find enough participants for my study who would agree to consume protein for several months.

Man: Oh, bad news!

Woman: Yeah, that could have been a good study but now it looks like I'll have to redesign it.

Man: So what exactly do you want to redesign? Like, change the topic or something?

Woman: Uh, in fact I can't really change the topic since I have already done a lot for the theoretical part of my thesis. I guess I'll just need to redesign my study in such a way that it won't require too many participants but I don't really see how I can do that.

Man: Hm, so your main problem is to find the participants who would agree to eat protein, right?

Woman: You got it.

Man: And how many participants do you need?

Woman: At least 15, and now I have only 10 people who contacted me and agreed to participate.

Man: Well, I suppose I can help you.

Woman: Really? How?

Man: You see, I am a weightlifter. I go to the rec center three times a week to work out so I won't mind getting some protein to boost my muscle growth.

Woman: Wow, I didn't know you are weightlifter! So you say you could volunteer to participate in my study?

Man: Sure. In addition, I have a bunch of friends there who lift weights with me. I am going to see them tonight so I'll talk to them about your study.

Woman: That would be awesome!

Man: I'm pretty sure most of them won't mind joining me. So I believe you will have enough people for your study.

Woman: Oh, Mark, thank you so much! I really appreciate your help.

Man: No problem, Mary. I will give you a call tonight and we'll discuss everything in detail.

Woman: OK. I will talk to you later then.

Man: Bye!

Now you will hear 5 questions. After you listen to each question, choose the best answer. You will have 12 seconds to answer each question.

Question 1: According to the conversation, which of the following is true about Mary's problem?

A. Mary can't find enough participants for her research study.

B. Mary doesn't want to graduate in spring.

C. Mary is too busy to work on her thesis.

D. Mary doesn't want to eat protein for several months.

Question 2: What is Mary planning to investigate in her study?

A. Why people eat protein when they want to grow muscles.

B. How much protein a person can eat within several months.

C. How protein affects muscle growth.

D. Why people do not want to eat protein every day.

Question 3: Why does Mark decide to help Mary?

A. He can redesign Mary's study.

B. He wants to eat some protein to increase his muscles.

C. There is no one else who agreed to participate in Mary's study.

D. He has more friends in the rec center than Mary does.

Question 4: According to the conversation, what can be inferred about protein?

A. All people like to eat protein.

B. Protein might help muscles grow.

C. It is difficult to eat protein for a long period of time.

D. Protein is not good for people who work out at the rec center.

Question 5: Each of the following was mentioned in the conversation EXCEPT

A. Mark will ask his friends to participate in Mary's project.

B. Mark will give Mary a call after he talks to his friends.

C. Mary didn't know that Mark is a weightlifter.

D. Mark goes to work out in the rec center every day.

This is the end of Listening Passage 3.

Listening Passage 4

Listen to the beginning of a lecture in a linguistics class.

Good morning. Today I am going to talk about standardization of a language. I am sure you've all heard such expressions as "standard language" and "standard English"… But

what exactly is “standard”? Why do we call some language variety “standard”? How is it different from a “nonstandard” language?

Well, first let me give you a definition of “standardization”. Standardization is basically a codification of a language, or arranging it in a systematic order. It is a process that involves the development of such things as grammar books, dictionaries, and literature in general. Standardization of a language also involves the agreement about what language rules and norms should be used. So, a standard language is a variety of a language that is used in print and usually taught in schools.

Now, how does a language become standardized? Well, there are several ways language standardization can happen. In some cases, standardization occurs as a result of a long-term process of language polishing, refinement, and enrichment by certain literary figures, such as Shakespeare. In other cases, language standardization is a political process that involves governments or official bodies such a Ministry or Department of Education that decide what should be in the language and what shouldn’t be. In any case, it is an ideological matter because the standardization of a language symbolizes some kind of unification and solidarity among its speakers, giving them prestige and, uh… sometimes even power over those who do not use the standard form.

Okay, so now let’s talk about the differences between standard vs. nonstandard language, but before that I’d like to emphasize that there is no a clear-cut borderline between them. So, standard language can have both written and spoken forms and can be used in both formal and informal situations. Nonstandard language, in its turn, includes colloquial language and slang. Although colloquial language is understood and accepted by many people and most of colloquial expressions can even be found in dictionaries, they are

considered to be inappropriate in formal settings. Nevertheless - and I find it quite interesting - colloquial expressions can eventually evolve and become part of the standard language. Uh… and as far as slang goes, it's the least formal and the most nonstandard variety of a language that is used by some speakers in informal situations. And slang is usually a temporary phenomenon, which means that slang expressions often become out-of-date after a while.

Finally, I want you to think about standard and nonstandard not as two completely separate language varieties but as a continuum from most to least formal because, as I've already mentioned, there's no clear-cut distinction between them.

Now you will hear 5 questions. After you listen to each question, choose the best answer. You will have 12 seconds to answer each question.

Question 1: According to the lecture, which of the following best describes language standardization?

A. Codification of a language according to certain language rules and norms.

B. Process of teaching language at schools.

C. Agreement about which literature should be used for reading.

D. Unification of speakers by the government.

Question 2: Each of the following was mentioned about the process of language standardization EXCEPT

A. Language standardization can be a long-term process.

B. A government can be involved in the process of language standardization.

C. Standardization of a language creates solidarity among its speakers.

D. Standardization of a language can be approved only by a Ministry or Department of Education.

Question 3: Why does the woman mention Shakespeare?

A. To show how well he could use standard language.

B. To give an example of a literary figure who influenced language standardization.

C. To suggest that Shakespeare knew both standard and non-standard language.

D. To draw attention to his writing skills.

Question 4: According to the lecturer, what can be inferred about people who use non-standard language?

A. They may not have as much power and prestige as people who speak the standard language.

B. They never studied the standard language at school.

C. They don't know who Shakespeare was.

D. They think that the non-standard language is easier than the standard language.

Question 5: According to the lecture, which of the following is true about colloquial expressions?

A. They always turn into slang.

B. They are never included in dictionaries.

C. They can become part of the standard language.

D. They are considered appropriate in formal settings.

This is the end of Listening Passage 4.

Listening Passage 5

Listen to the beginning of a lecture in a journalism class.

In today's class we will be discussing different media in which journalism as a discipline exists. As you know, journalism applies to a number of different media such as newspapers, magazines, radio, TV, and the Internet. But to make things simple, we can split journalism into three main branches: print journalism, broadcast journalism, and online journalism.

So let's start with print journalism first. It includes newspapers, magazines, and private publications. Journalists in this branch work at various tasks, for example, editing, reporting, photography, page layout, illustration and graphics. Print journalism is characterized by a written mode that requires brevity and presentation of most important facts. Because of the written mode, there is always a gap between the time when events take place and when they are reported in a newspaper of magazine. Think about it: it takes time to write a newspaper article and print it before it gets to a reader. Other distinguishing characteristics of print journalism are storage, accessibility, and portability – you don't need to have any technology for reading a newspaper, and you can easily carry it with you and get back to reading it anywhere and anytime.

Uh, the next branch of journalism is broadcast journalism that includes radio and television. Unlike print journalism, broadcast journalism relies on auditory and, in the case of TV, visual information. Thus, broadcast journalists must not only gather and present important information, but also record interesting and relevant video and sounds. In order to do that, they need special technical skills for video and audio recording and editing. Both radio and TV journalists usually do not have as much time between the events and the

presentation of information as print journalists do. In addition, broadcast journalism requires technical support – you won't be able to hear or see news unless you have a radio or a TV set.

And finally, online journalism and a couple words about its unique nature. Although being relatively new, online journalism possesses a number of features that make if different from print and broadcast journalism. These features include interactivity, multimedia, and hypertext. Hyperlinks can be used for easy navigation, multimedia complements news stories with video, audio, and graphics, and interactivity allows for communication with other web-users. Oh, and here's the interesting part: the World Wide Web provides potentially unlimited storage capacity for online journalism and allows journalists to disseminate news almost immediately. And we can only guess how the constant technological development will shape this type of journalism. Needless to say, online journalism requires special skills and expertise in graphic design, multimedia design, site production, and programming.

So, even though different media have different impacts on journalism, they all are irreplaceable and have their place with respect to different audiences and different needs.

Now you will hear 5 questions. After you listen to each question, choose the best answer. You will have 12 seconds to answer each question.

Question 1: What makes broadcast journalism different from print journalism?

A. Use of visuals.

B. Use of interactivity.

C. Use of sound and video.

D. Use of graphic design.

Question 2: According to the lecture, which of the following is NOT true about print journalism?

A. Print journalism exists only in a written form.

B. There is always time between the events and the presentation of information.

C. One of the main characteristics of print journalism is its accessibility.

D. Print journalism is more popular than broadcast journalism and online journalism.

Question 3: The professor mentioned all of the following about broadcast journalism EXCEPT that:

A. It is easily accessible for all people.

B. It relies mostly on the use of audio and video.

C. It requires special technical skills.

D. It involves audio and video editing.

Question 4: According to the lecture, what can be inferred about the three types of journalism?

A. Print journalism is the most popular type of journalism.

B. Out of three types of journalism, online journalism is likely to change most.

C. Broadcast journalism developed much faster than online journalism.

D. Broadcast journalism presents only the most important information.

Question 5: Why does the professor say that all types of media are irreplaceable for journalism?

A. She wants to make clear that there are no better or worse types of media for journalism.

B. She intends to emphasize the advantages and disadvantages of media.

C. She thinks that journalism and media mean the same thing.

D. She believes that there are too many different types of media that need to be replaced.

This is the end of Listening Passage 5.

Listening Passage 6

Listen to a conversation on campus between two students.

Man: Professor Johnson's lecture was pretty interesting today, don't you think? I particularly enjoyed the part where he was talking about the mysteries of Ancient Egypt.

Woman: Well, Jim, some of that stuff was hard to believe though… I mean how could they build those pyramids without any modern machinery?

Man: What do you mean, Megan?

Woman: Well, Professor Johnson said that the stone blocks they used in building the pyramids weighed more than 2.5 tons with some of the blocks being as heavy as 15 tons! Can you imagine that? I mean, whoever built those pyramids, how could they move those blocks without any machines? No way!

Man: You're right, Megan! That seems difficult to believe. But remember what Professor Johnson said – there is still no real consensus among scientists about pyramid construction techniques as well as what kind of workforce was used, yet the pyramids are there, built as they are.

Woman: Hm… And how about the measurements of the Great Pyramid of Giza? According to what Professor Johnson said, the base of the Great Pyramid is flat and horizontal to 15 mm, and its sides are aligned to the four cardinal compass points. Also, the stones are placed

in such an exact contact that it's impossible to insert a tip of a knife between them. I don't believe the ancient Egyptians could have built the pyramids with such a precision…

Man: Uh, why? Who do you think built the pyramids then?

Woman: Well, the other day, I watched a TV program and I remember they were saying that in those times such structures as the Great Pyramids of Giza could hardly be constructed by humans.

Man: What are you talking about? You are not saying that some aliens came from space and built the pyramids for the pharaohs, are you?

Woman: In fact, that was the theory they were discussing on TV. Yeah, Jim, I know it sounds kind of weird but don't you think it could be possible?

Man: Of course, and computers were also brought to us from Mars.

Woman: Don't make fun of me! I am just telling you what I saw on TV.

Man: Yeah, well OK Megan, I have to get over to the library. I still have to finish my paper for Mechanical Engineering class and it's due tomorrow.

Woman: Is it that big group assignment that you have been working on for several weeks?

Man: No, it's a small research project. I just need to finish the conclusion part.

Woman: Well, good luck! I'll see you then in class on Thursday.

Man: Thanks! See you later!

Now you will hear 5 questions. After you listen to each question, choose the best answer. You will have 12 seconds to answer each question.

Question 1: What was Megan's attitude to Professor Johnson's discussing the mysteries of Ancient Egypt?

A. She believed everything Professor Johnson said in class.

B. She found Professor Johnson's discussion very interesting.

C. She doubted that the pyramids had been built by the ancient Egyptians.

D. She was surprised that Professor Johnson didn't know how the pyramids had been built.

Question 2: According to the conversation, which of the following was NOT mentioned by Professor Johnson about the pyramids?

A. The pyramids were built using the blocks that weighed from 2.5 to 15 tons each.

B. The measurements of the pyramids are very precise.

C. Pyramid construction techniques used by the ancient Egyptians remain unknown.

D. There is evidence that the pyramids were built using machinery.

Question 3: Listen again to part of the conversation. Then answer the following question.

Man: What are you talking about? You are not saying that some aliens came from space and built the pyramids for the pharaohs, are you?

Woman: In fact, that was the proposed theory on TV. Yeah, Jim, I know it sounds kind of weird but don't you think it could be possible?

Man: Of course, and computers were also brought to us from Mars.

What can be inferred about Jim's attitude toward Megan's theory?

A. He agreed with Megan's explanation.

B. He thought that computers were brought from Mars.

C. He didn't believe that the pyramids had been built by the aliens.

D. He was surprised that Megan watched TV.

Question 4: Why was Jim going to the library?

A. To finish the conclusion part for his small research project.

B. To work on a big group project.

C. To get some books for his Mechanical Engineering class.

D. To find information about the pyramids.

Question 5: According to the passage, which of the following is true about Jim and Megan?

A. They are dating.

B. They are in the same class.

C. They are relatives.

D. They met for the first time today.

This is the end of Listening Passage 6.

Congratulations!

You just finished the Listening Test!

Please fill out the Post-Test Questionnaire.

Thank you for your time and effort!

APPENDIX B. SCREEN SHOTS OF AN ONLINE VERSION OF THE LT

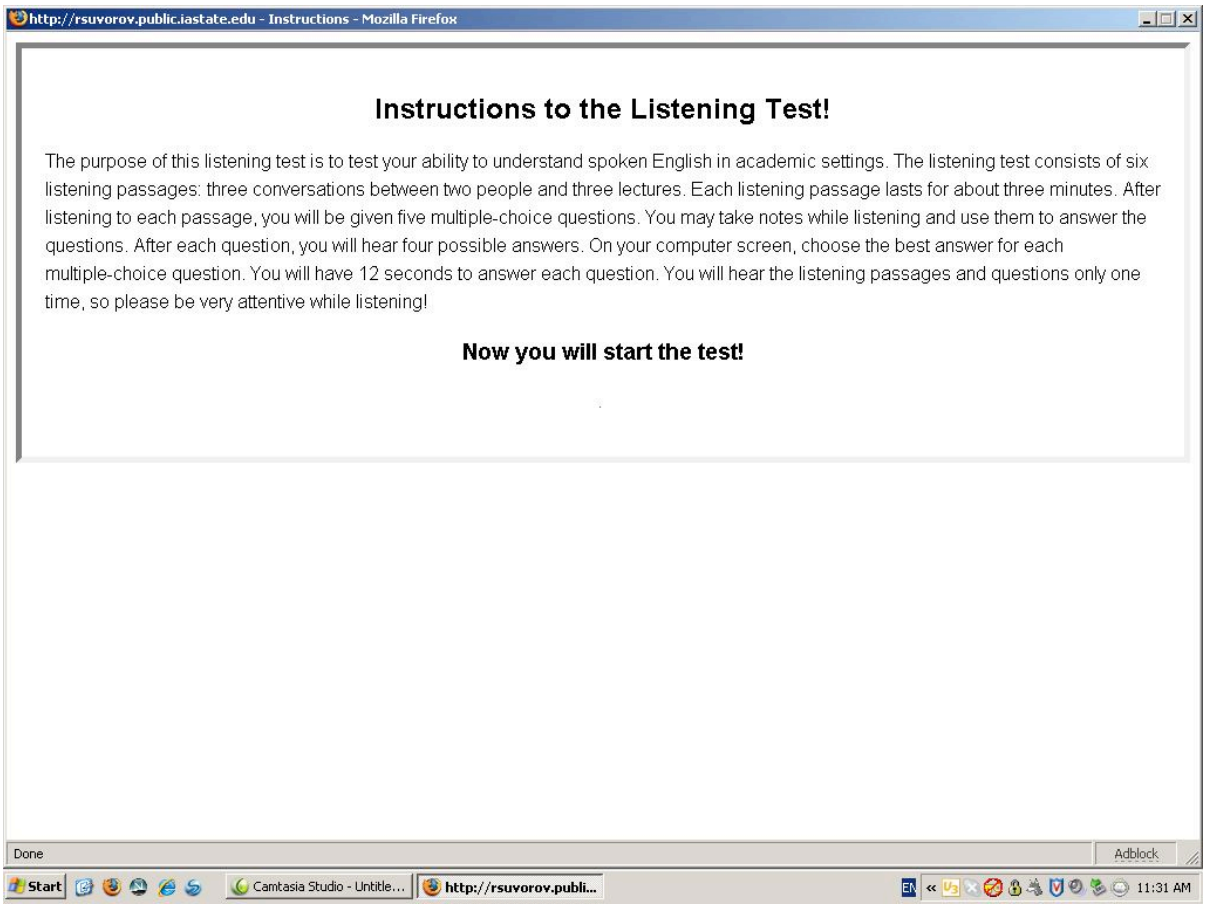

Figure B1. Instructions to the Listening Test

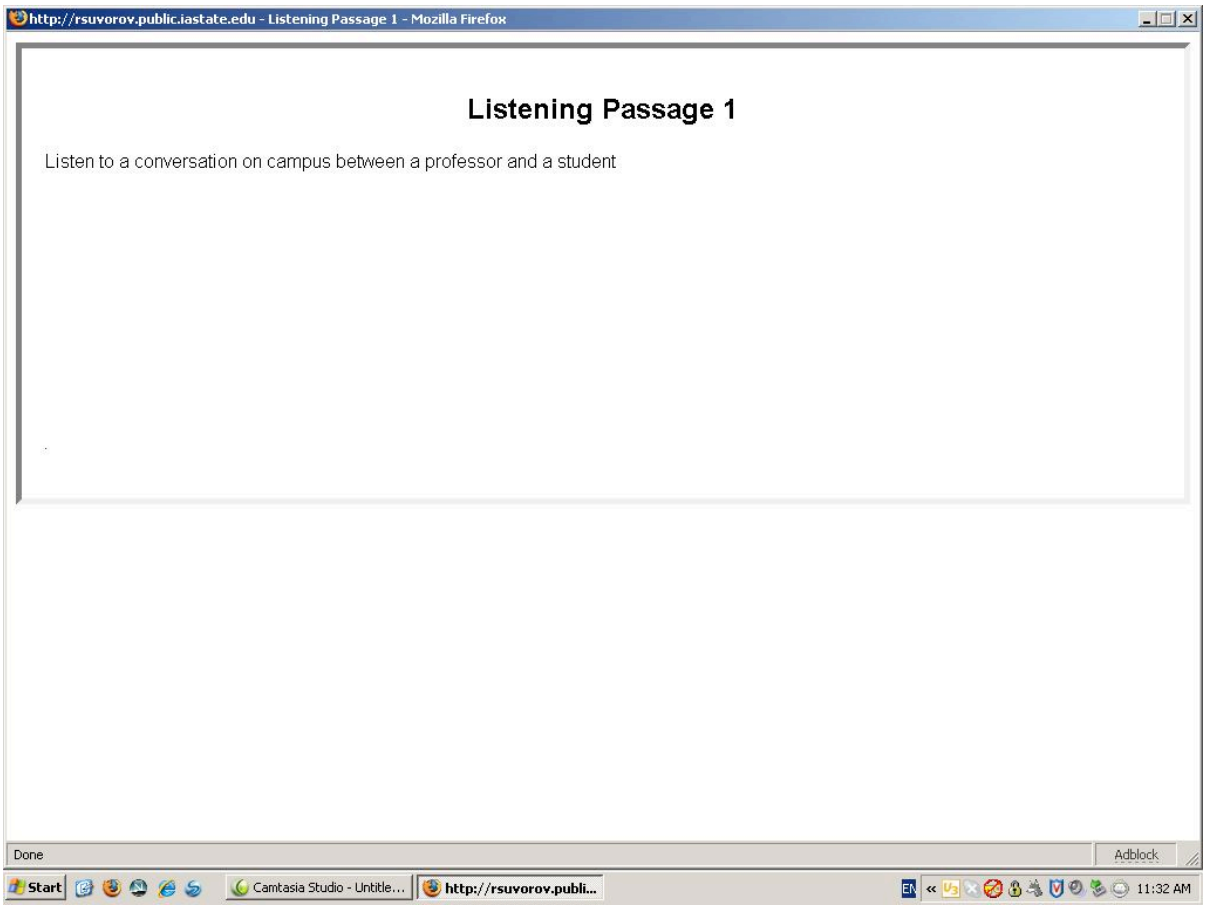

Figure B2. Audio-only listening passage

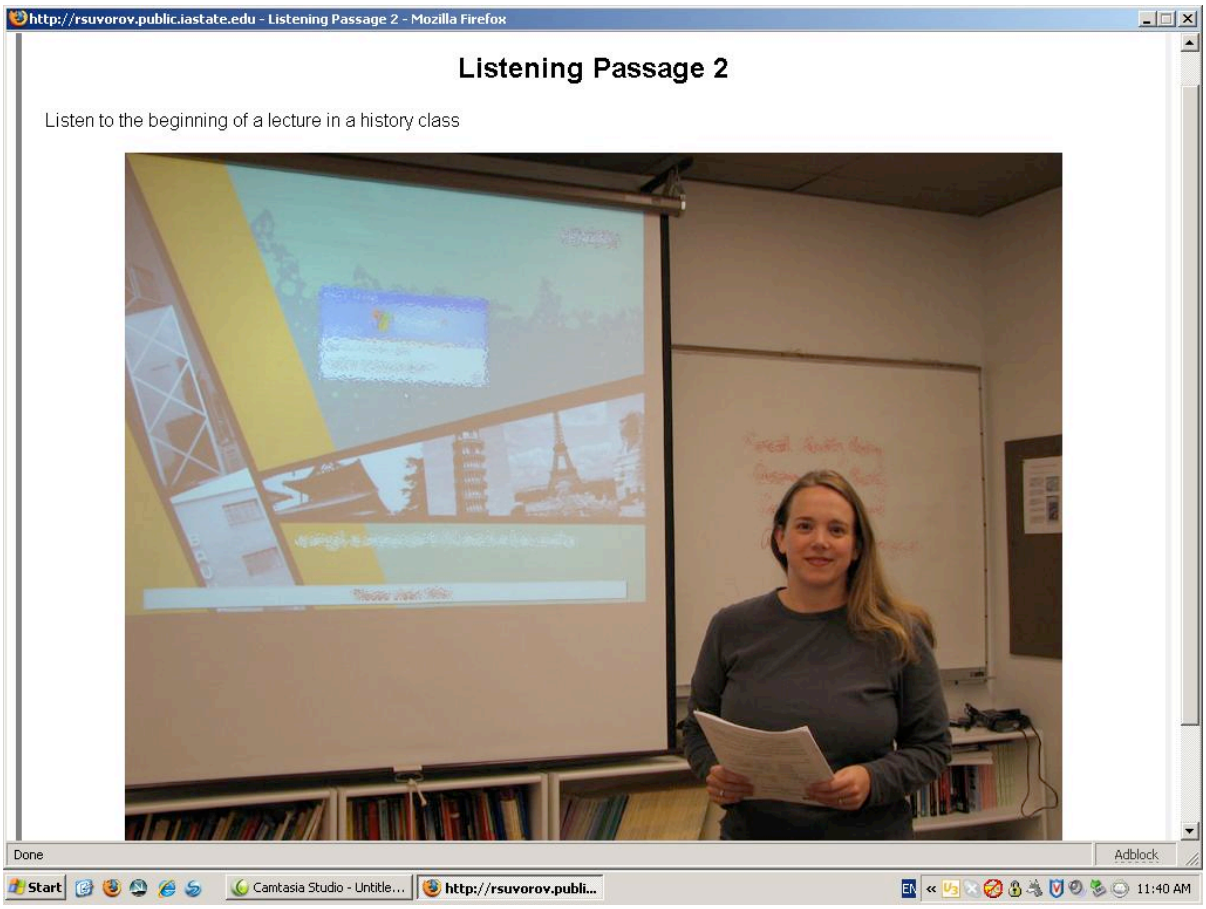

Figure B3. Listening passage with a photograph

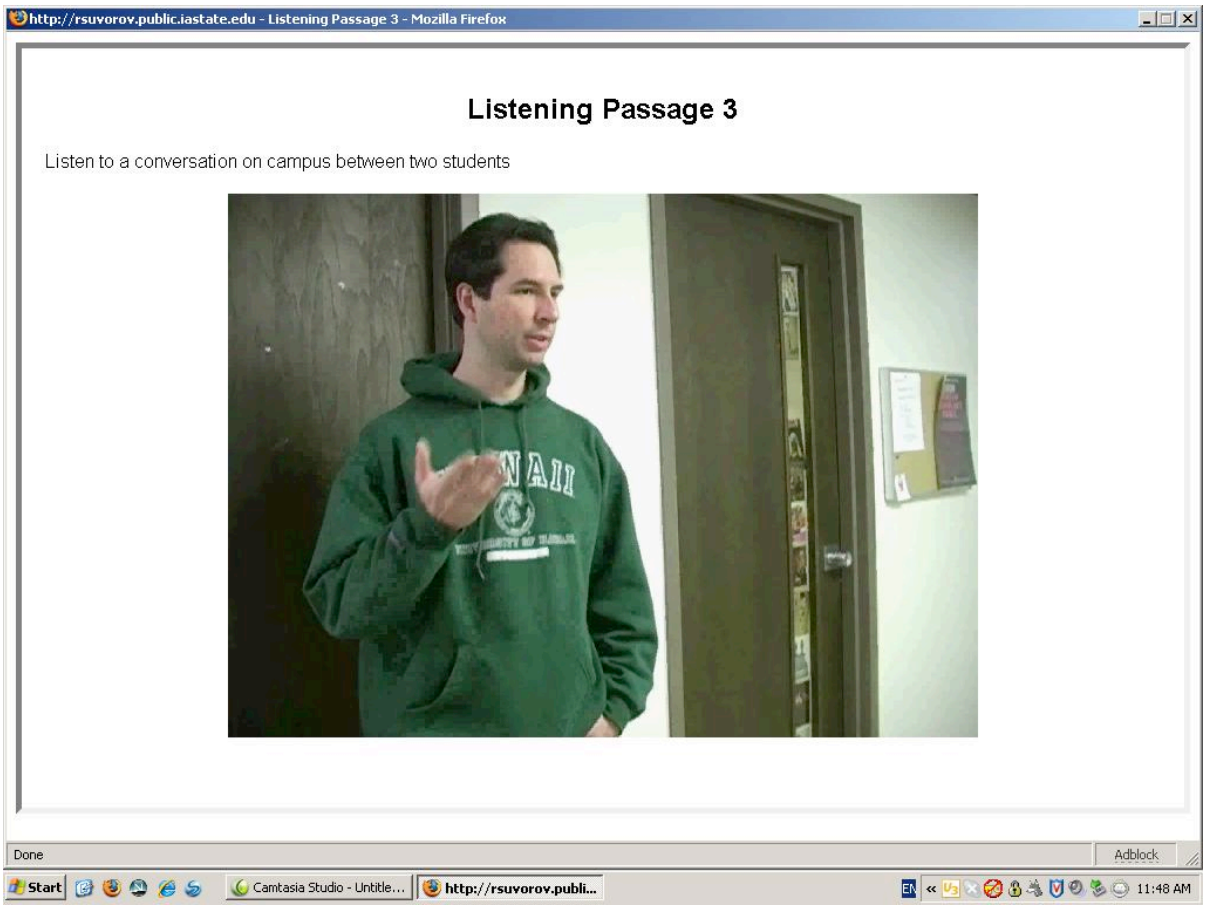

Figure B4. Video-mediated listening passage

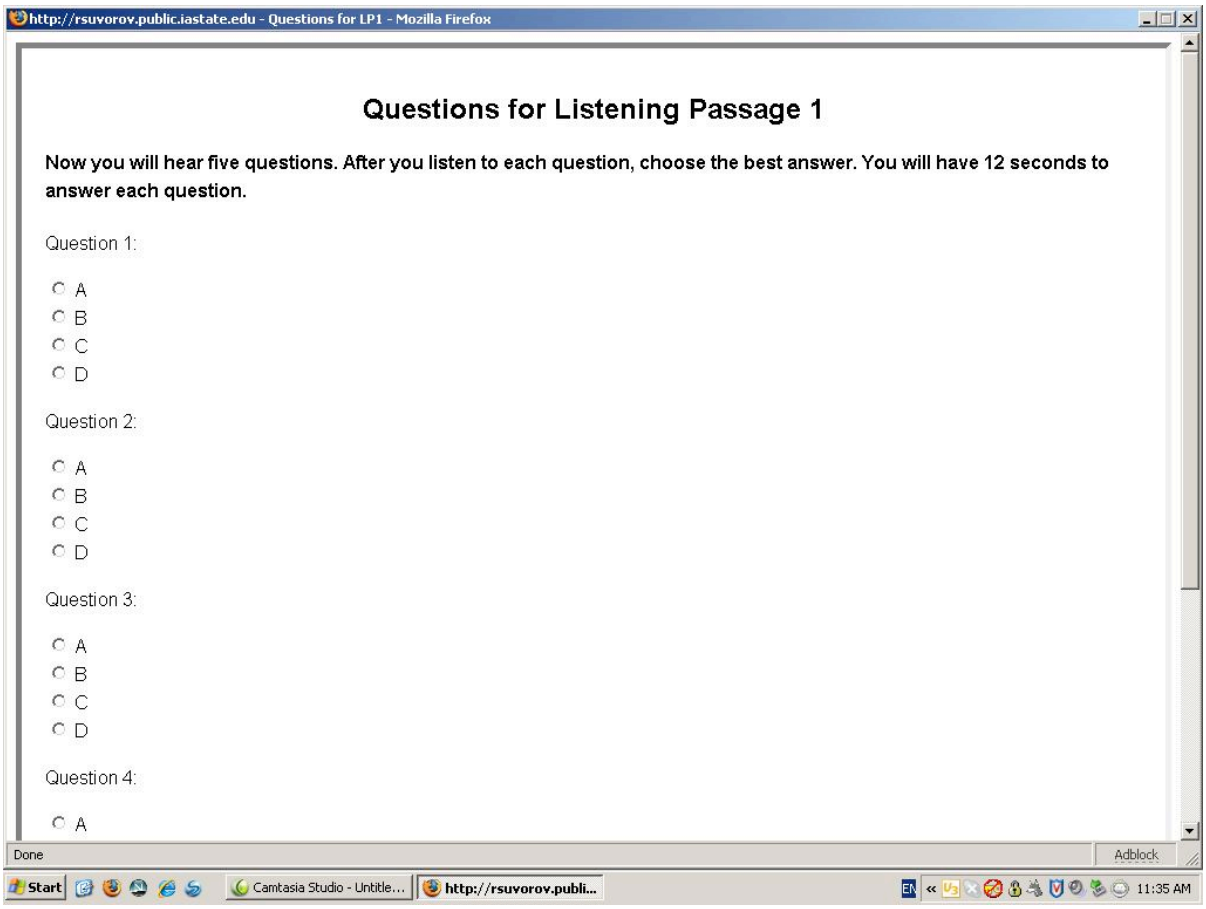

Figure B5. Questions for a listening passage

APPENDIX C. TABLE OF SPECIFICATIONS

Table C1

Table of Specifications

Types of questions	LP1	LP2	LP3	LP4	LP5	LP6	Total number	%
True-false (T)	Q2	Q2	Q1	Q5	Q2	Q5	6	20
Exception (E)	Q5	Q4	Q5	Q2	Q3	Q2	6	20
Inference (I)	Q3	Q1	Q4	Q4	Q4	Q3	6	20
Details (D)	Q4	Q3	Q2	Q1	Q1	Q1	6	20
Purpose (P)	Q1	Q5	Q3	Q3	Q5	Q4	6	20
Total number	5	5	5	5	5	5	30	100

APPENDIX D. PRE-TEST QUESTIONNAIRE

1. Name (first, last): _____
2. Age: _____
3. Gender: Female ____ Male ____
4. Country of origin: _____
5. Native language(s): _____
6. How many years have you been studying English: _____
7. How long did you live in the US: ____ years _____ months
8. Which class are you currently enrolled in? 99L ___ IEOP High Listening ___
9. Your major field of study at ISU (if any): _____
10. Do you watch TV here in the US? Yes ____ No ____

    If YES, how often do you watch TV in the US:

    Less than once a week ___ Once a week ___ 2-3 times a week ___ Every day ___
11. Do you have problems understanding the English language when you watch movies in English?

    Never ___ Rarely ___ Sometimes ___ Most of the time ___ All the time ___
12. Do you have problems when you talk on the phone with native speakers of English?

    Never ___ Rarely ___ Sometimes ___ Most of the time ___ All the time ___
13. How often did you use video (e.g. TV news, movies, or video courses) for learning English?

    Never ___ Rarely ___ Sometimes ___ Most of the time ___ All the time ___
14. How often did you use audio (e.g. radio, CD recordings, or audio tapes) for learning English?

    Never ___ Rarely ___ Sometimes ___ Most of the time ___ All the time ___

Thank you for your answers!

APPENDIX E. PERCEPTUAL LEARNING PREFERENCES SURVEY

(Adapted from Kinsella, 1993)

DIRECTIONS: The purpose of this survey is to understand the ways you prefer to learn. Read each question and check (√) the answer that most accurately describes how you learn.

	USUALLY	SOMETIMES	RARELY
1. I can remember most of the information I have heard in a lecture or class discussion without taking notes.			
2. I learn more by reading about a topic than by listening to a lecture or a class discussion.			
3. When I study new material, I learn more easily by looking over visual aids (such as charts and illustrations) in a chapter than by reading the assigned pages.			
4. Talking about a subject with someone else helps me better understand my own ideas.			
5. I take notes during class lectures and discussions and read them carefully several times before a test.			
6. When I read a textbook, newspaper, or novel, I picture the ideas or story in my mind.			
7. I remember information that I have discussed in class with a partner or a small group better than information that I have read or written about.			
8. I get confused when I try to figure out graphs and charts that do not come with a written explanation.			
9. I remember information well by listening to tapes.			
10. To remember a new word, I must hear it and say it.			
11. I would rather see a film on a subject than listen to a lecture or read a book or magazine article.			
12. I prefer reading a newspaper or magazine as a source of news rather than listening to the radio or watching the television.			
13. I make drawings in my study notes or on study cards to remember new vocabulary and important material.			

	USUALLY	SOMETIMES	RARELY
14. I read assigned material and notes aloud to myself to concentrate and understand better.			
15. When I listen to an explanation or lecture, I form mental images or pictures to understand better.			
16. I best understand homework or test instructions by reading them on the board or on a handout rather than by just listening to them.			
17. It is easier for me to remember illustrations and charts in textbooks if they are done in bright colors.			
18. I prefer to watch the television or listen to the radio for news rather than to read a newspaper or a magazine.			
19. To remember a new word, I must see it several times.			
20. Before making or drawing something, I first picture in my mind what my completed project will look like.			
21. I find it difficult to figure out what to do on homework assignments when the teacher just gives us a handout without discussing it in class.			
22. I have difficulty understanding a new term if I have only a definition with no examples or illustrations.			
23. I regularly read newspapers, magazines, or books for pleasure and information.			
24. When I have homework reading assignments, I take notes or summarize the main ideas in writing.			

Thank you for your answers!

APPENDIX F. POST-TEST QUESTIONNAIRE

1. Name (first, last): _____
2. Overall, how would you describe the difficulty level of the listening tests:

   Very easy ___ Easy ___ Normal ___ Difficult ___ Very difficult ___
3. Did you have problems understanding the speakers? Yes ___ No ___

   If yes, please explain:
   _____
   _____
4. Was some information from the lectures familiar to you? Yes ___ No___

   If yes, please give details:
   _____
   _____
5. Did you take notes while listening?

   Never ___ Rarely ___ Sometimes ___ Most of the time ___ All the time ___
6. If you took notes, did they help you answer the questions after listening? Yes ___ No ___

   Please explain:
   _____
   _____
7. How often did you look at the computer screen when listening?

   Never ___ Rarely ___ Sometimes ___ Most of the time ___ All the time ___
8. Did the **video** help you better understand the speakers? Yes ___ No ___

   Please explain:
   _____
   _____
9. Did the **pictures** help you better understand the speakers? Yes ___ No ___

   Please explain:
   _____
   _____
10. Did the quality of audio and/or video affect your understanding of the speakers?

    Please explain:
    _____
    _____

-> Go to the NEXT PAGE -->

11. Did you prefer the **video** version of the listening tests or the **pictures** version? Why?

    _____
    _____

12. Did the visual information (i.e. pictures or video) ever distract you from listening? Yes __ No __

    If yes, please explain:

    _____
    _____

13. What was the hardest thing about the listening tests?

    _____
    _____

14. If you were to choose between the three types of the listening tests (i.e. video version, picture version, and audio-only version), which one would you prefer?

    _____
    _____
    _____

15. Any other comments about the listening tests?

    _____
    _____

Thank you for your answers!

APPENDIX G. TEST-TAKERS' SCORES ON THE LISTENING TEST

Table G1

Test-Takers' Scores on the Listening Test

Student	Group	AD	PL	VD	AL	VL	PD	Total
1	IEOP	4	2	4	3	1	1	15
2	IEOP	3	5	4	4	5	4	25
3	IEOP	4	5	4	4	5	5	27
4	IEOP	2	4	2	1	3	3	15
5	IEOP	3	1	3	4	2	2	15
6	IEOP	3	3	4	2	2	5	19
7	IEOP	2	3	3	4	2	5	19
8	IEOP	4	2	3	5	2	5	21
9	IEOP	3	4	3	3	3	5	21
10	IEOP	4	4	4	4	2	5	23
11	IEOP	4	5	4	2	1	5	21
12	IEOP	4	2	3	2	2	4	17
13	99L1	5	2	3	2	1	3	16
14	99L1	4	4	4	3	2	4	21
15	99L1	5	3	4	4	3	3	22
16	99L1	1	2	4	2	1	1	11
17	99L1	5	5	4	3	2	4	23
18	99L1	4	1	2	4	2	4	17
19	99L1	4	2	3	5	4	2	20
20	99L1	3	2	3	2	1	1	12
21	99L1	4	2	4	1	1	3	15
22	99L1	4	3	2	2	1	5	17
23	99L1	0	3	4	3	2	3	15
24	99L1	2	1	1	3	1	1	9
25	99L1	4	2	3	2	2	5	18
26	99L2	4	4	5	2	1	3	19
27	99L2	4	4	3	4	1	3	19
28	99L2	4	4	4	5	2	4	23
29	99L2	5	3	3	2	3	4	20
30	99L2	3	3	4	1	2	2	15
31	99L2	3	3	1	1	0	5	13
32	99L2	5	2	3	5	0	4	19
33	99L2	1	3	2	1	1	1	9
34	99L2	5	1	2	2	0	2	12

Note. A – audio, P – photograph, V – video, D – dialogue, L – lecture.

APPENDIX H. TEST-TAKERS' RESPONSES ON POST-TEST QUESTIONNAIRE

Table H1

Test-Takers' Responses on Post-Test Questionnaire

Student	Group	[a]Difficulty	[b]Note-taking	[c]Notes usefulness	[c]Video usefulness	[c]Photo usefulness	[c]Visuals distraction	[d]Format preferred
1	IEOP	4	3	0	0	0	1	A
2	IEOP	3	5	0	0	0	0	V
3	IEOP	3	3	1	0	0	0	A
4	IEOP	4	4	1	1	0	0	V
5	IEOP	4	3	1	1	0	1	V
6	IEOP	3	2	0	0	1	1	P
7	IEOP	4	3	1	1	0	0	V
8	IEOP	3	5	1	0	0	1	A
9	IEOP	3	4	1	0	0	1	A
10	IEOP	3	4	1	1	0	1	V
11	IEOP	3	4	1	0	0	0	V
12	IEOP	3	2	1	0	0	1	P
13	99L1	3	2	1	0	0	0	A
14	99L1	4	1	0	0	0	0	V
15	99L1	3	4	1	0	0	0	P
16	99L1	4	3	1	1	1	0	V
17	99L1	2	3	1	0	0	0	V
18	99L1	3	3	1	0	0	0	A
19	99L1	3	3	1	0	0	0	V
20	99L1	3	3	1	0	0	1	P
21	99L1	3	3	1	0	1	1	P
22	99L1	4	4	1	0	0	1	P
23	99L1	4	4	1	1	0	0	A
24	99L1	4	3	0	0	0	0	A
25	99L1	3	4	1	1	0	1	A
26	99L2	2	1	-	0	0	0	A
27	99L2	3	4	1	0	0	0	A
28	99L2	3	4	1	0	0	0	A
29	99L2	2	4	1	1	0	0	V
30	99L2	4	5	1	0	1	0	P
31	99L2	3	3	1	0	0	0	A
32	99L2	3	3	1	0	0	1	A
33	99L2	4	5	1	1	1	0	V
34	99L2	4	4	1	0	0	0	A

Note. [a]Difficulty level of the Listening Test: 1 – very easy, 2 – easy, 3 – normal, 4 – difficult, 5 – very difficult. [b]Frequency of note-taking: 1 – never, 2 – rarely, 3 – sometimes, 4 – most

of the time, 5 – all the time. [c]Helpfulness of note-taking, usefulness of video, usefulness of photographs, and distraction by visuals: 1 – yes, 2 – no. [d]Preferred format of a listening passage: A – audio-only, V – video-mediated, P – photo-mediated.

REFERENCES

Baltova, I. (1994). The impact of video on the comprehension skills of core French students. *Canadian Modern Language Review, 50*(3), 507-531.

Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). *TOEFL 2000 listening framework: A working paper* . Princeton, NJ: Educational Testing Service.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System, 25*(1), 39-53.

Brindley, G. (1998). Assessing listening abilities. *Annual Review of Applied Linguistic, 18,* 171-191.

Buck, G. (2001). *Assessing listening.* Cambridge: Cambridge University Press.

Chang, A. C-S., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(2), 375-397.

Chung, U. K. (1994). *The effect of audio, a single picture, multiple pictures, or video on second-language listening comprehension.* Unpublished PhD dissertation, University of Illinois at Urbana-Champaign.

Coakley, C. G., & Wolvin, A. D. (1986). Listening in the native language. In B. H. Wing (Ed.), *Listening, reading, and writing: Analysis and application. Northeast Conference Reports: Northeast Conference on the Teaching of Foreign Languages* (pp. 11-42). Middlebury, VT: The Conference.

Coniam, D. (1999). Subjects' reactions to computer-based tests. *Journal of Educational Technology Systems, 27*(3), 195-206.

Coniam, D. (2001). The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *System, 29*, 1-14.

Ginther, A. (2001). *Effects of the presence and absence of visuals on performance on TOEFL CBT listening-comprehension stimuli* (Report 66). Princeton, New Jersey: Educational Testing Service.

Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing, 19*(2), 133-167.

Gruba, P. (1993). A comparison study of audio and video in language testing. *JALT Journal, 15*(1), 85-88.

Gruba, P. (1997). The role of video media in listening assessment. *System, 25* (3), 335-345.

Gruba, P. (1999). *The role of digital video media in second language listening comprehension.* Unpublished PhD dissertation, Department of Linguistics and Applied Linguistics, University of Melbourne. Retrieved May 5, 2008, from http://eprints.unimelb.edu.au/archive/00000244/

Gruba, P. (2006). Playing the videotext: A media literacy perspective on video-mediated L2 listening. *Language Learning and Technology, 10* (2), 77-92.

Hearst, M. A. (2000). The debate on automated essay grading. *IEEE Intelligent Systems*, *15*(5), 22-37.

Jones, L. C., & Plass, J. L. (2002). Supporting listening, comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal, 86*(4), 546-561.

Jones, L. C. (2003). Supporting listening comprehension and vocabulary acquisition with multimedia annotations: The students' voice. *CALICO Journal, 21*(1), 41-65.

Kinsella, K. (1995). Perceptual learning preferences survey. In J. M. Reid (Ed.), *Learning styles in the ESL/EFL classroom* (pp. 221-238). Boston, MA: Heinle & Heinle.

Lado, R. (1961). *Language testing: The construction and use of foreign language tests.* London: Longman.

Lynch, T. (1998). Theoretical perspectives on listening. *Annual Review of Applied Linguistics, 18,* 3-19.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.

Mendelsohn, D.J., & Rubin, J. (1995). *A guide for the teaching of second language listening.* San Diego, CA: Dominie Press.

Mueller, G. A. (1980). Visual contextual cues and listening comprehension: An experiment. *Modern Language Journal, 64*(3), 335-340.

Ockey, G. J. (2007). Construct implications of including still image or video in computer-based listening tests. *Language Testing, 24*(4), 517-537.

Paivio, A. (1986). *Mental representations: A dual-coding approach.* Oxford, England: Oxford University Press.

Parry, T., & Meredith, R. (1984). Videotape vs. audiotape for listening comprehension tests: An experiment. (ERIC Document Reproduction Services ED 254 107).

Progosh, D. (1996). Using video for listening assessment: Opinions of test-takers. *TESL Canada Journal, 14*(1), 34-44.

Rubin, J. (1994). A review of second language listening comprehension research. *Modern Language Journal, 78*(2), 199-221.

Rubin, J. (1995). The contribution of video to the development of competence in listening. In D. Mendelsohn, & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 151-165). San Diego, CA: Dominie Press.

Salomon, G. (1989). 4 learning from texts and pictures: Reflections on a meta-level. In H. Mandel, & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures.* Amsterdam: Elsevier.

Schriver, K. A. (1997). *Dynamics in document design: Creating text for readers.* New York, NY: John Wiley & Sons, Inc.

Secules, T., Herron, C., & Tomasello, M. (1992). The effect of video context on foreign language learning. *The Modern Language Journal, 76*(4), 480-490.

Thompson, I. (1995). Assessment of second/foreign language listening comprehension. In D. Mendelsohn, & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 31-58). San Diego, CA: Dominie.

Wagner, E. (2007). Are they watching? Test-taker viewing behavior during an L2 video listening test. *Language Learning and Technology , 11*(1), 67-86.

Wetzel, C. D., Radtke, P. H., & Stern, H. W. (1994). *Instructional effectiveness of video media.* Hillsdale, NJ: Lawrence Erlbaum.

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist, 11*(2), 87-95.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*(4), 345-376.

Zenisky, A. L., & Sireci, S. G. (2002). Technological innovations in large-scale assessment. *Applied Measurement in Education, 15*(4), 337-362.

ACKNOWLEDGEMENTS

It would not have been possible to write this thesis without the help and support of many people whose contributions I would like to acknowledge here.

I would like to thank Erik Voss and Kimberly LeVelle, doctoral students in Applied Linguistics and Technology at Iowa State University, who have contributed their time and effort to help with audio and video recording of the listening passages. I am very grateful to Adolfo Carrillo Cabello and Tommy Lage for their assistance with administering the Listening Test to all three groups of participants. I would also like to thank statistician Man-Yu Yum for her help with the statistical analysis of my data.

In addition, I am greatly indebted to my major professor, Dr. Volker Hegelheimer, for his enthusiastic support, invaluable advice and generous help throughout my work on this research study, and to my thesis committee members, Dr. Dan Douglas and Dr. Charles Kostelnick, for their helpful feedback and expertise on the topic.

Finally, special thanks go to my dear friend Réka for her unwavering inspiration, warm encouragement, and emotional support throughout this year, as well as her significant help with my pilot study.